Epigenetic characterisation of *ANK1*, a novel gene implicated in Alzheimer's disease (AD)

Submitted by

Adam Robert Smith

To the University of Exeter

as a thesis for the degree of

Doctor of Philosophy in Medical Studies

in March 2018

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature……  ……

ABSTRACT

Alzheimer's disease is a progressive neurodegenerative disorder that affects approximately twenty eight million people worldwide. Collectively DNA mutations only explain approximately 33% of Alzheimer's disease incidence. Therefore, Alzheimer's disease has been hypothesised to have both an environmental and epigenetic contribution to disease aetiology. Epigenetics refers to the heritable changes in gene expression that do not involve changes to the underlying DNA sequence. Epigenetic changes can be transient, varying in levels according to a person's age, sex, environmental exposure, genetics and disease status. The identification of robust epigenetic changes in Alzheimer's disease would give the potential to open up a variety of new treatment options for the disease.

Previous epigenome-wide association studies in Alzheimer's disease have highlighted neuropathology-associated DNA hypermethylation in the *ANK1* gene. To date, however, no studies have examined whether other epigenetic mechanisms are altered in this gene in Alzheimer's disease pathology, or whether these *ANK1* epigenetic changes are seen in other neurodegenerative diseases.

The aim of this thesis was to characterise the epigenetic profile (DNA methylation, DNA hydroxymethylation and histone modifications) of *ANK1* in Alzheimer's disease. In addition, it sought to determine whether the previously identified *ANK1* DNA methylation changes in Alzheimer's disease are specific to this dementia, or are observed in a number of other neurodegenerative diseases. The results provide further support for a role for epigenetic dysfunction of *ANK1* in Alzheimer's disease as well as some other neurodegenerative diseases.

In summary, the work presented in this thesis represents the first epigenome-wide association study of DNA methylation and DNA hydroxymethylation simultaneously in Alzheimer's disease across two brain regions, highlighting changes in DNA modifications in *ANK1*. The thesis then highlights *ANK1* DNA methylation changes in other neurodegenerative diseases and explores the histone modification profile of the *ANK1* locus in Alzheimer's disease.

ACKNOWLEDGEMENTS

If anyone tells you that “doing a PhD is easy” then they’re lying. However, my PhD has been made considerably easier by the amazing help and guidance I have received along the way.

I would like to start by thanking my first supervisor, Professor Katie Lunnon for all of her hard work, not only finding funding for my PhD project in the first place, but understanding, guiding and sharing your expertise and enthusiasm with me throughout the entire project. I have heard her described as an inspiration for female researchers, whereas in truth, she is an inspiration for all researchers. I would also like to thank my second supervisor, Professor Jonathan Mill for being so supportive, kind and always managing to find time to discuss an issue, despite his busy schedule.

I would like to thank everyone within the Complex Disease Epigenetics Group, especially those that make my day to day job so enjoyable. Bex, for all her guidance, support and laughter she was effectively my third supervisor and will always be my first point of call for all things R and cat related. Ehsan, my green brother, I’m sure our time as the Mario brothers will go down in the team’s history, although I’m keeping my full beard from now on. Joe, thank you for all your support in the lab I have learned so much from you. Isabel, always full of drive and enthusiasm, I hope that this approach to research has brushed off on me. Janou and Jenny, have always been on hand for me to complain to, laugh with/at, despair with or have stupid conversations with. They have both managed to keep me sane (sort of), help with the decorating of my desk (cats, elephants, snowmen, smurfs, the TARDIS and frowns!) and obviously help me celebrate now this is all over. Thank you also to all the collaborators, students and staff of Exeter University who helped me in any way during my PhD. Thank you all and I hope our work together continues for many years to come. I also want to thank BRACE a charity that do brilliant work raising funds for dementia research, including my own, thank you for making all this possible.

I would like to thank my family for being the most fun/mad family in the world. I want to especially thank my parents for all the opportunities, encouragement

and love you have given me. But most importantly, thank you for being the most amazing parents ever.

Finally, I would like to dedicate this work, along with any future achievements to my own little family. My greatest thanks and deepest love goes to my wife Alix, I would not have achieved what I have without your love and support. Every day you are the one that inspires me to do better and push that bit harder, it's been a long three and a half years but I hope I have made you proud. You, Arla and our little one on the way, are the reason I got through my PhD and I love you all so very much.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

PUBLICATIONS ARISING FROM THIS THESIS

Chapter 1 (published manuscript presented in Appendix 1):

Smith, A. R.; Mill, J.; Smith, R. G.; Lunnon K. (2016). Elucidating novel dysfunctional pathways in Alzheimer's disease by integrating loci identified in genetic and epigenetic studies. Neuroepigenetics 6: 32-50.

Chapter 2 (submitted manuscript presented in Appendix 2):

Smith, A. R.; Lunnon, K. (2018). Post-mortem brain tissue and its application to study epigenetic regulation in Alzheimer's disease. Post Mortem Neurochemistry. V. Balcar. Springer Nature. Under Review

Chapter 3 (submitted manuscript presented in Appendix 3):

Smith, A. R.; Smith, Rebecca G; Hannon, Eilis; Roubroeks, Janou A Y; Burrage, Joe; Troakes, Claire; Al-Sarraj, Safa; Mill, Jonathan; van den Hove, Daniel L; Lunnon Katie (2018). Parallel profiling of DNA methylation and hydroxymethylation highlights neuropathology-associated epigenetic variation within *ANK1* and thirteen novel loci in Alzheimer's disease entorhinal cortex. Under Review

Chapter 4 (submitted manuscript presented in Appendix 4):

Smith, A. R.; Smith, Rebecca G; Burrage, Joe; Troakes, Claire; Al-Sarraj, Safa; Kalaria, Rajesh N; Sloan, Carolyn; Robinson, Andrew C ; Mill, Jonathan; Lunnon, Katie (2018). A cross-brain-regions study of *ANK1* DNA methylation in different neurodegenerative diseases. Under Review

During the course of my PhD I have also co-authored these other publications:

Crawford, Bethany; Craig, Zoe; Mansell, Georgina; White, Isobel; **Smith, Adam R.**; Spaull, Steve; Imm, Jennifer; Hannon, Eilis; Wood, Andy; Yaghootkar, Hanieh; Ji, Yingjie; Frayling, Tim; Mullins, Niamh; Lewis, Cathyrn; Mill, Jonathan; Murphy, Therese M. (2018). "DNA methylation and inflammation marker profiles associated with a history of depression." Under Review

Hannon, E., E. Dempster, J. Viana, J. Burrage, **A. R. Smith**, R. Macdonald, D. St Clair, C. Mustard, G. Breen, S. Therman, J. Kaprio, T. Toulopoulou, H. E. Hulshoff Pol, M. M. Bohlken, R. S. Kahn, I. Nenadic, C. M. Hultman, R. M. Murray, D. A. Collier, N. Bass, H. Gurling, A. McQuillin, L. Schalkwyk and J. Mill (2016). An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation. Genome Biol 17(1): 176.

Marzi, S., T. Ribarska, **A. R. Smith**, E. Hannon, J. Poschmann, K. Moore, C. Troakes, S. Al-Sarraj, S. Beck, S. Newman, K. Lunnon, L. Schalkwyk and J. Mill (2017). A histone acetylome-wide association study of Alzheimer's disease: neuropathology-associated regulatory variation in the human entorhinal cortex. bioRxiv.

Smith, A R., R. G. Smith, D. Condliffe, E. Hannon, L. Schalkwyk, J. Mill and K. Lunnon (2016). "Increased DNA methylation near TREM2 is consistently seen in the superior temporal gyrus in Alzheimer's disease brain." Neurobiol Aging 47: 35-40.

Smith, R. G. H., E; De Jager, PL; Chibnik, L; Lott, SJ; **Smith, A R.**; Condliffe, D; Lunnon, K (2018). Elevated DNA methylation across a 48kb region spanning the *HOXA* gene cluster on chromosome 7 is associated with Alzheimer's disease neuropathology in the prefrontal cortex and superior temporal gyrus. Alzheimer's & Dementia

Wright, P. B., Joe; **Smith, Adam R.**; Hamilton, Patrick; Morris, Colin; Dempster, Emma; Mathews, Fiona; Schofield, Henry (2018). Application of a novel molecular method to age free-living wild Bechstein's bats. Under Review

DECLARATIONS

The London Neurodegenerative Diseases Brain Bank (LNDBB) provided the samples used in Chapters 3 and 5. The samples used in Chapter 4 were obtained from six of the UK Brain Banks: Cambridge Brain Bank, LNDBB, Manchester Brain Bank, Newcastle brain bank, Oxford Brain Bank and the South West Dementia Brain Bank (SWDBB).

In Chapters 3-5, all laboratory work was carried out by myself under the guidance of Dr Joe Burrage and Dr Katie Lunnon, with one exception: in Chapter 3, Oxidative bisulfite conversions were performed with the assistance of Miss Janou Roubroeks, a PhD student within the group.

All bioinformatic and statistical analysis performed throughout Chapters 3-5 was carried out by me, under the guidance of Dr Rebecca Smith and Dr Katie Lunnon.

ABBREVIATIONS

Abbreviation	Term
Δ	Corrected DNA methylation difference
450K array	Illumina Infinium 450K Methylation BeadChip Array
5caC	5-carboxylcytosine
5fC	5- formylcytosine
5hmC	5-hydroxymethylcytosine
5hmU	5-hydroxymethyluracil
5mC	5-methylcytosine
AD	Alzheimer's disease
ALS	Amyotrophic lateral sclerosis
BER	Base excision repair
BP	Biological process
BS	Bisulfite
CAA	cerebral amyloid angiopathy
CC	Cellular component
CER	Cerebellum
CETS	Cell epigenotype specific
CGIs	CpG Islands
ChIP	Chromatin Immunoprecipitation
ChIP-qPCR	Chromatin Immunoprecipitation quantitative PCR
chr	Chromosome
CME	Clathrin-mediated endocytosis
CSF	Cerebrospinal fluid
DHP(s)	Differentially hydroxymethylated position(s)
DHR(s)	Differentially hydroxymethylated region(s)
DLB	Dementia with Lewy bodies
DMoP(s)	Differentially modified position(s)
DMP(s)	Differentially methylated position(s)
DMR(s)	Differentially methylated region(s)
DNA	Deoxyribose nucleic acid
DNMT	DNA methyltransferase
dNTP(s)	Deoxynucleotide(s)
DP	Distal promoter
DS	Downstream
EC	Entorhinal cortex
EOAD	Early onset Alzheimer's disease
EPIC array	Illumina Infinium EPIC Methylation BeadChip Array
ESC	Embryonic stem cells
EWAS	Epigenome-wide association study
FAD	Familial Alzheimer's disease
GB	Gene body
GO	Gene ontology
GWAS	Genome-wide association study

H3K27me3	Tri-methylation of lysine at position 27 of histone 3
H3K4me3	Tri-methylation of lysine at position 4 of histone 3
HBSS	Hanks’ balanced salt solution
HCP(s)	High CpG content promoter(s)
HD	Huntington's disease
hMeDIP	Hydroxymethylated DNA immunoprecipitation
HS	Hereditary Spherocytosis
IG	Intergenic
LCP(s)	Low CpG content promoters(s)
LD	Linkage disequilibrium
LNDBB	London Neurodegenerative Diseases Brain Bank
LOAD	Late onset Alzheimer’s disease
MAF	Minor allele frequency
MAP	Microtubule-associated protein
MCI	Mild cognitive impairment
MeDIP	Methylated DNA immunoprecipitation
MF	Molecular function
miRNA(s)	Micro-RNA(s)
NC	Outside CpG island
NFT(s)	Neurofibrillary Tau tangle(s)
NTC	No template control
OxBS	Oxidative Bisulfite
PBS	Phosphate buffered saline
PCI	Phenol / Chloroform / Isoamyl Alcohol
PCR	Polymerase chain reaction
PD	Parkinson's disease
PET	positron-emission tomography
PGS	Polygenic risk scores
PHFs	Paired helical filaments
PIC	Protease inhibitor cocktail
PP	Proximal promoter
QC	Quality control
qPCR	Quantitative polymerase chain reaction
RNA	Ribose nucleic acid
RRHP	Reduced representation hydroxymethylation profiling
SHE	CpG island shelf
SHO	CpG island shore
SN	Substantia nigra
SNP	Single nucleotide polymorphism
STG	Superior temporal gyrus
STR	Striatum
SWDBB	South West Dementia Brain Bank
TBE	Tris-borate EDTA
TdT	Terminal deoxynucleotidyl transferase
TET	Ten-eleven translocation dioxygenase

TPM	Transcripts per million
TSS	Transcription start sites
UA	Unannotated
VaD	Vascular dementia
vLDL	Very low density lipoproteins
β-amyloid	Amyloid-beta

CHAPTER 1: INTRODUCTION

This chapter introduces research that has been performed in the field of Alzheimer's disease (AD) and epigenetics to date. The work presented in this chapter comparing genetic and epigenetic studies in AD is based on a published review (Smith et al. 2016), which can be found in Appendix 1.

1.1. Alzheimer's disease

Dementia is an umbrella term that encompasses a large number of neurological diseases, including AD, Dementia with Lewy bodies (DLB), Vascular dementia (VaD), Huntington's disease (HD) and Parkinson's disease (PD) dementia. Each disease has phenotypic symptoms or pathological hallmarks that are specific to that disease; however they all result in the deterioration of the mental ability of patients leading to a demented state. AD is a progressive neurodegenerative disorder characterised by the accumulation of amyloid-beta (β-amyloid) plaques and neurofibrillary Tau tangles (NFTs) within the brain (Querfurth et al. 2010) and accounts for approximately 60% of the dementia prevalence around the world (Fratiglioni et al. 1999).

1.1.1. Types of AD

AD is sub-categorised into early onset (EOAD), which occurs before the age of 65 and late onset (LOAD), which occurs after 65 years of age. EOAD accounts for a very small percentage (1-5%) of total AD cases and is usually caused by autosomal dominant mutations in one of three genes (Amyloid precursor protein (*APP*), Presenilin1 (*PSEN1*) and Presenilin2 (*PSEN2*)) (Bekris et al. 2010). As a result this form of AD is directly heritable and is termed familial AD (FAD). Although AD is typically thought to be a disease of old age, EOAD can present considerably earlier, in patients as young as 40.

LOAD is far more common, accounting for ~95% of AD incidence. Apart from the difference in age of onset, the characteristic pathological hallmarks of these two classes of AD are very similar. However, one study has shown significant differences in the Praxis, which tests the ability to correctly imitate hand gestures and voluntarily mime tool use and Temporal Orientation performance, which is related to a tendency for impairment in visual memory of EOAD

compared to LOAD patients. This suggests a localised disease sub-type specific development, with EOAD patients having a greater left posterior hemisphere susceptibility and LOAD having a more localized disease to the limbic structures (Sá et al. 2012).

1.1.2. Prevalence and economic burden

An estimated 28 million people around the world are currently living with AD, with this figure predicted to continue rising (Prince et al. 2015); the number of people living with AD in Western Europe alone is projected to increase from 5 million in 2015 to over 9.5 million by 2050 (Prince et al. 2015). In the UK specifically, over 560,000 people are currently living with AD (Prince et al. 2014). In 2015, the leading cause of death for women and the second leading cause for men in the UK was AD and other dementias, accounting for 15% and 8% of deaths, respectively (McLaren 2015). This has put a growing strain on the health care systems in the UK and around the world, with 42% of unplanned UK hospital admissions, being from people over the age of 70 suffering from dementia (Sampson et al. 2009).

Coupled to this high demand for care comes a substantial cost; an estimated £13.8 billion is spent on AD care each year in the UK. This is comprised of health care costs of £800 million, social care costs of £5.5 billion and the greatest cost being the £7.5 billion value for the time given by unpaid carers of AD patients. Due to an increasingly ageing population this is predicted to rise to £39.3 billion by 2050 (Lewis et al. 2014). This pattern is reflected globally, with the cost of AD projected to reach $0.6 trillion dollars by 2018 and will double to $1.2 trillion by 2030 (Wimo et al. 2017).

1.1.3. Disease progression

The progression of AD is typically slow and is comprised of three general stages — mild (early-stage), moderate (middle-stage) and severe (late-stage). Specific AD symptoms vary between patients and occur at different times in the disease progression. Despite this, the most common earliest symptoms are a worsening ability to retain new information leading to disruptions in daily routine.

This is followed by a difficulty in planning/solving problems, confusion related to time and/or place, speech trouble and mood/personality changes (Burns et al. 2009). Although the progression of disease is patient-specific the late stages tend to result in a lack of patient movement (Brunnström et al. 2009). At that point there is an increased risk of infections such as aspiration pneumonia, which is the leading cause of death in AD patients (Leonard 2011). As a result of the differing symptom severity between patients, being able to distinguish different stages in a clinical setting is challenging, this is made even more complex by inaccessibility of the brain for diagnostic testing and the delay between pathological and symptomatic onset. By the time individuals become symptomatic, there is already considerable neuronal cell loss, plaque deposition, and tangle burden within the brain, which can appear up to 10 years before a clinical diagnosis is made (Jack et al., 2010). As such, the only reliable method of AD diagnosis is a post-mortem measure of NFT pathology spread throughout the brain, called Braak staging (Braak et al. 1991). Braak stages I/II are used when neurofibrillary tangle pathology is confined mainly to the transentorhinal region of the brain, stages III/IV when there is also involvement of limbic regions, and V/VI when there is extensive neocortical involvement (Braak et al. 1991). As such, the earliest site of AD pathology is the entorhinal cortex (EC) which exhibits increased pathology as Braak staging progresses. The cerebellum has the lowest levels of NFT pathology even at the very last stages of disease (Braak et al. 1989). Both EC and CER were assessed in this thesis, as they represent high and low levels of pathology respectively.

1.1.4. Causes of AD

Although the neuropathology in AD has now been well characterised, the exact mechanisms driving disease onset and progression are still poorly understood. The focus of research is slowly shifting to drug targets and treatment strategies; however the vast majority of current AD research is still primarily focussed on aetiology, investigating molecular causes of disease. Research to date has nominated three main hypotheses as to the cause of disease pathology: the amyloid cascade hypothesis, the tau tangle hypothesis and the inflammatory hypothesis.

1.1.4.1. Amyloid cascade hypothesis

The amyloid cascade hypothesis is built around the presence of extracellular Aβ plaques in the brain of AD patients. Intermembrane bound *APP* is cleaved by Beta (β) secretase and Gamma (γ) secretase leaving a long amino-acid strand (sAPP-β) and a short amino acid strand (β-amyloid). These soluble β-amyloid protein monomers aggregate to form dense insoluble plaques. These plaques in turn cause neuritic injury, neuronal dysfunction and cell death, leading to AD (Hardy et al. 1992) (Figure 1.1). There are however numerous reports that the soluble β-amyloid precursors of these plaques are the cause of pathology, as they show greater correlation with AD symptoms than plaque burden (Lue et al. 1999, Naslund et al. 2000). It is hypothesised that that the APOE protein binds β-amyloid directly, with the protein encoded by the ε4 allele having a greater affinity for β-amyloid than the corresponding ε3 or ε2 alleles (Strittmatter et al. 1993, Strittmatter et al. 1993, Sanan et al. 1994). However, this was disputed in 1994 when the reverse was shown to be true (LaDu et al. 1994). More recent work has shown that the level of APOE affinity for β-amyloid is dependent on the lipidation status of APOE (Carter 2005), Further details on APOE and hypotheses involving its role in AD can be found in Section 1.1.4.3.

1.1.1.1. Tau tangle hypothesis

The presence of NFTs of hyperphosphorylated tau in the brain correlates more closely with neurodegeneration and disease symptoms than β-amyloid plaques (Figure 1.2). β-amyloid plaque levels plateau at their highest amount before the earliest cognitive symptoms appear in patients (Jack et al. 2010). Tau pathology, on the other hand, is still rising as patients experience mild cognitive decline and plateaus when patients reach full dementia status (Jack et al. 2010). Tau is a microtubule-associated protein (MAP), which is involved in the polymerisation of tubules and maintenance of neuronal structure. Hyperphosphorylation of the tau protein causes the removal of tau from the microtubule structure leading to destabilisation and degradation. Phosphorylated tau aggregates and form paired helical filaments (PHFs) and further disrupts normal microtubule function, mainly in the organisation of

cytoplasm (Kosik et al. 1986). PHF deposition leads to NFTs, a primary pathological hallmark of AD (Figure 1.3). The tau protein is actively phosphorylated in pathology-free individuals and is a mechanism of regulating neuronal axon length. It is therefore thought that this phosphorylation must be tightly regulated by phosphatases such as tau protein phosphatase (Gotz et al. 2001). Decreased function of this class of enzyme may be one possible explanation for the accumulation of NFTs in AD. β-amyloid has been found to have an effect on the deposition of NFTs increasing levels five-fold in transgenic mouse models, further highlighting the multifactorial nature of AD pathology (Gotz et al. 2001).

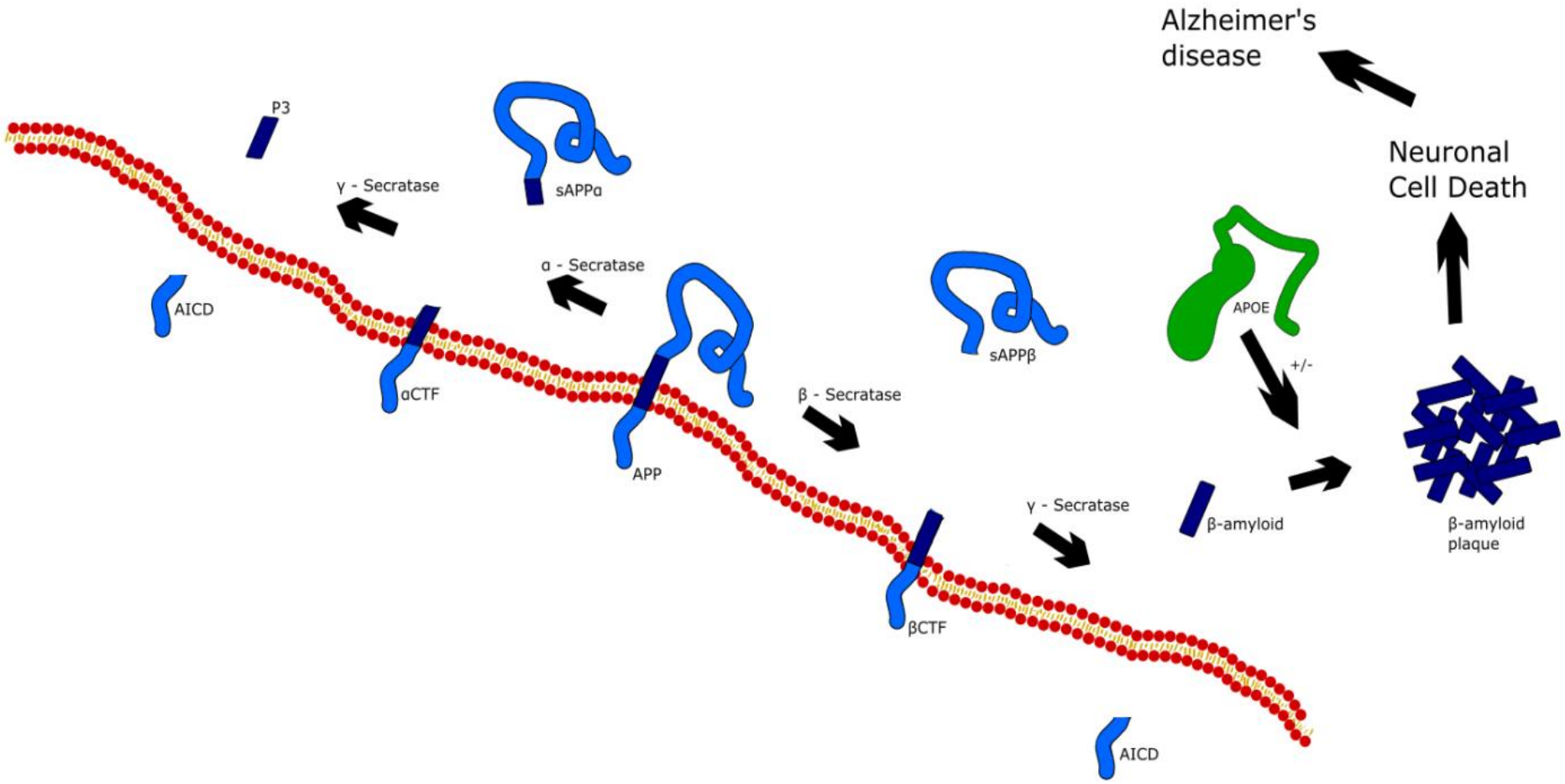


Figure 1.1; APP cleavage pathway by β and γ secretase leading to the accumulation of β-amyloid plaques typical of AD.

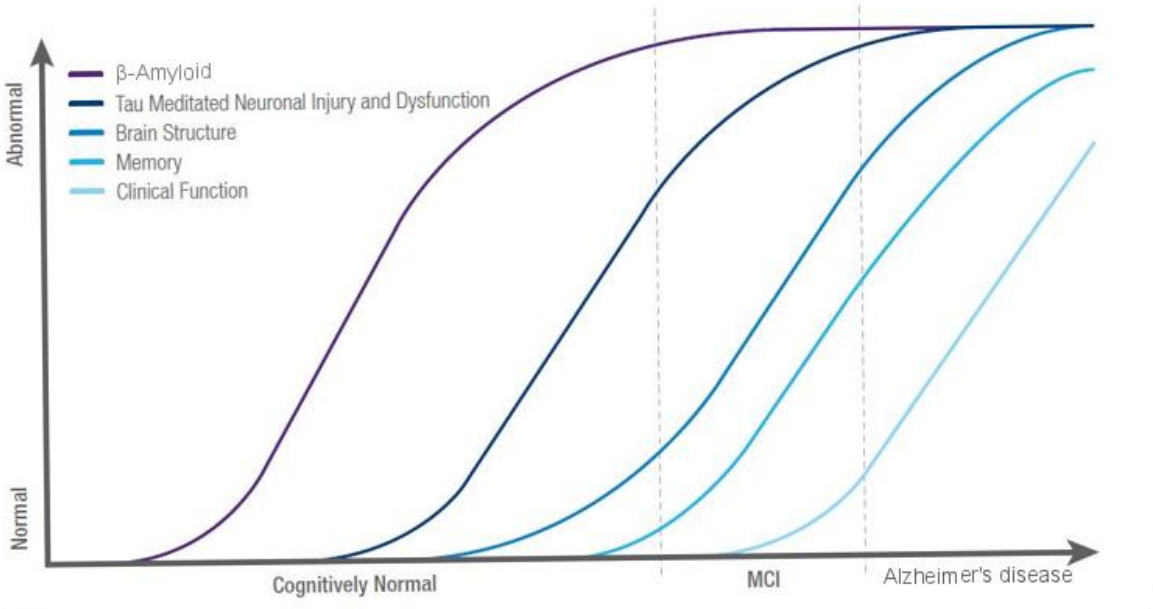


Figure 1.2; Graphical representation of the onset of pathological hallmarks of AD in relation to symptomatic onset*. (Taken from (Imagilys 2015))*

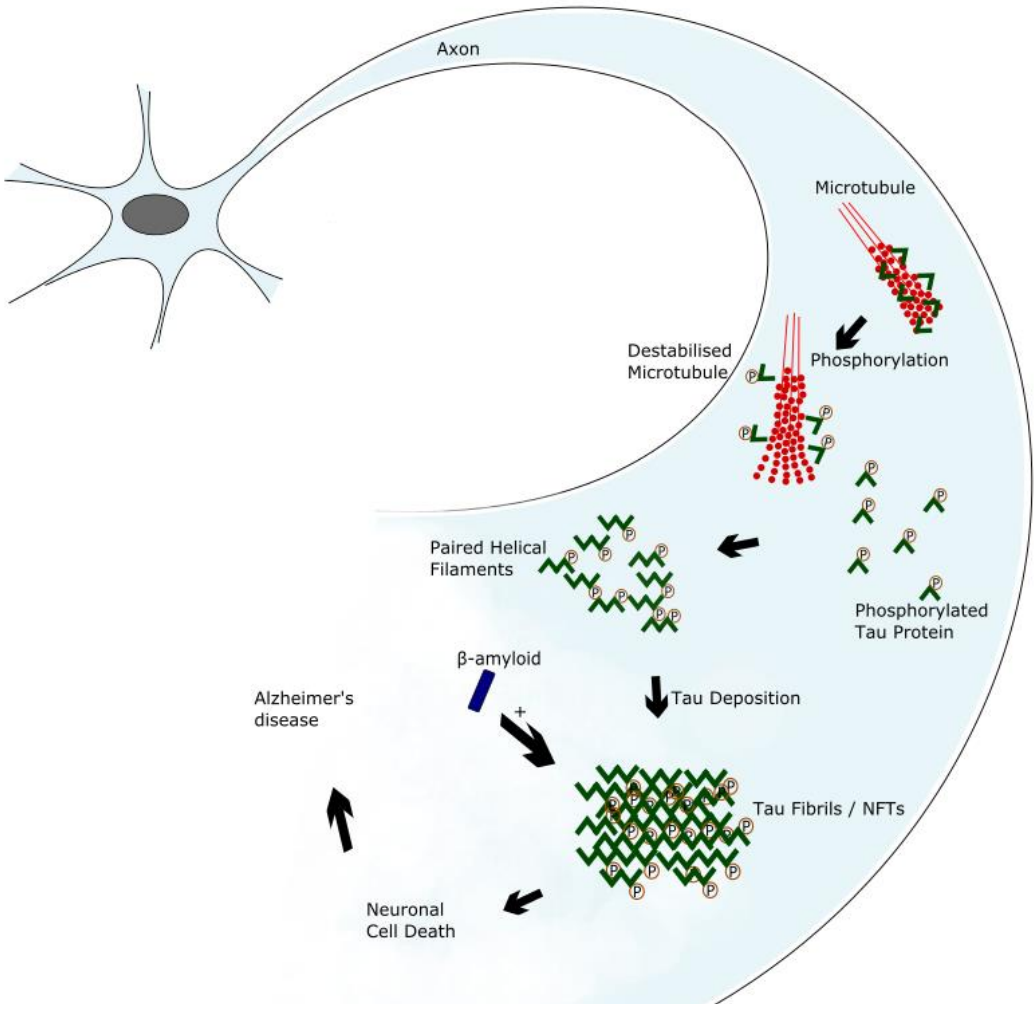


Figure 1.3; Schematic highlighting pathological changes in tau in AD. *Tau protein hyperphosphorylation leads to the formation of Paired Helical Filaments (PHFs) and ultimately Neurofibrillary Tangles (NFTs), one of the major hallmarks of AD.*

1.1.1.2. Inflammatory hypothesis

Both tau tangles and β-amyloid plaques are detrimental aggregates that occur in AD brain. These protein oligomers are usually cleared by microglial-regulated processes, before they have an impact on neuronal health and function (Daniel Lee et al. 2010). An impaired immune response could lead to errors in the clearance of these proteins in the brains of affected patients. Several studies have shown that the normal process of physiological ageing causes the pathological conformational changes of microglia (Norden et al. 2013), the immune regulators of the brain (Sheng et al. 1997), into a pre-activated state, termed priming. Increased expression of inflammatory markers including MHC class II and complement receptor 3 (CR3 - Cd11b/CD18) have been observed in the aged human and mouse brain (Streit et al. 2004, Godbout et al. 2005). In addition, evidence of activated microglial morphology has been identified in aged models of the brain; staining microglia for the cell surface protein Iba1 indicates that microglia from non-diseased healthy aged brains have shorter and less branched dendrites than those in younger animals (Choi et al. 2007, Hwang et al. 2008), morphology typical of primed microglia. Under standard homeostatic conditions, the function and abundance of microglia within the CNS is tightly regulated. However, in response to the aforementioned tau tangles and β-amyloid aggregates, primed microglia, which are hypersensitive to inflammatory stimuli, cause an exaggerated inflammatory response (Perry et al. 2014). The increased response resulting from this could, in turn, lead to further neuronal damage. This theory has been termed the inflammatory hypothesis (Figure 1.4).

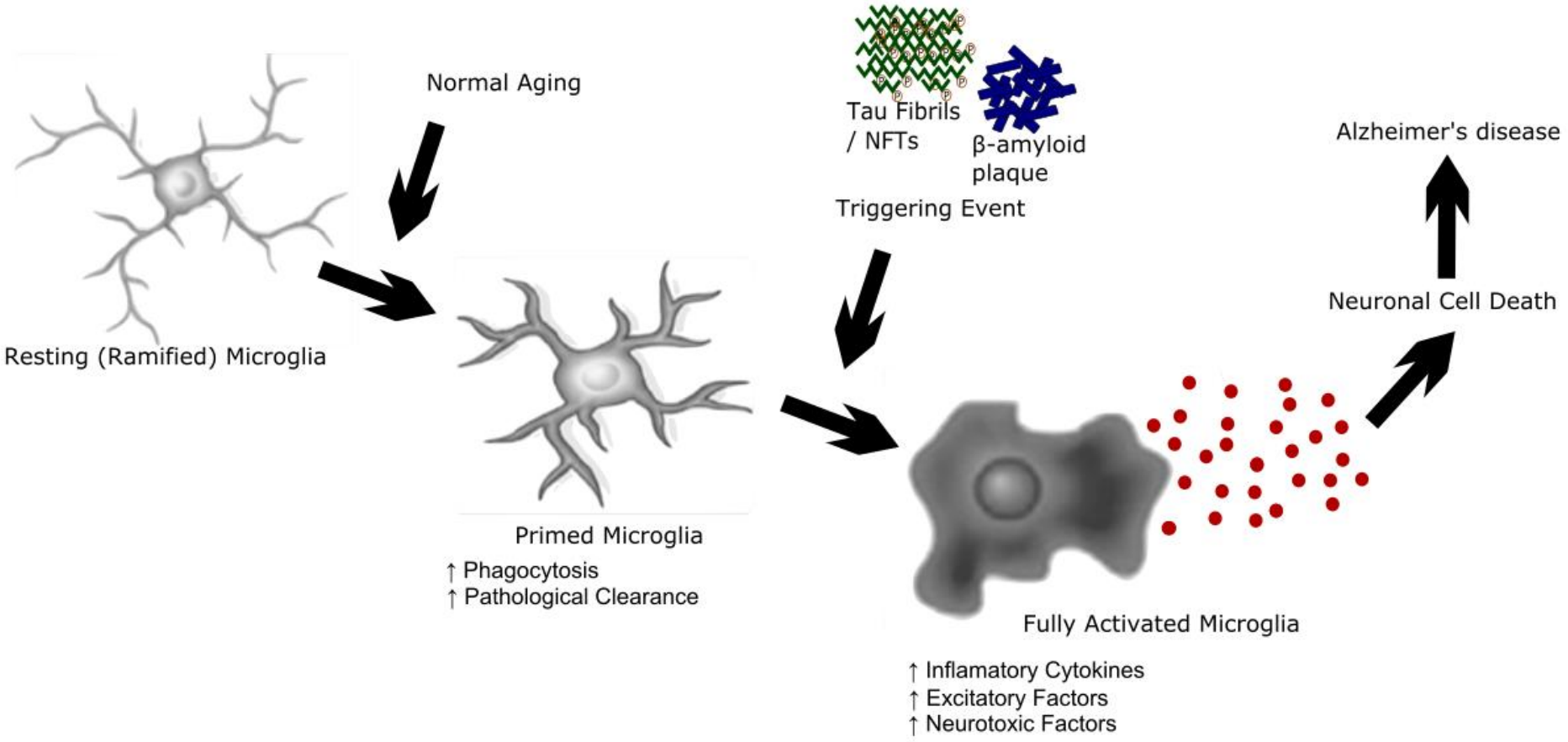


***Figure 1.4; Schematic showing the potential role of microglia in AD**. Resting microglia become primed by the normal process of ageing. With secondary insult (i.e. NFTs and β-amyloid plaques) these microglia go through a conformational change into a fully activated form, which have a neurotoxic phenotype. These microglia produce pro-inflammatory factors that cause damage to the neurones of an AD patient's brain.*

1.1.1.3. Genetic mechanisms

Quantitative genetic analyses have demonstrated high heritability estimates (58%-79%) for AD (Gatz et al. 2006), and thus initial approaches to understanding aetiology have focused on uncovering a genetic contribution to disease susceptibility. In recent years, the recruitment of large cohorts and the now relatively inexpensive cost of assessing genetic variation through genome-wide association studies (GWAS) has allowed the identification of multiple variants associated with an elevated risk of developing AD. Many of these genes have also been robustly associated with AD via subsequent meta-analyses (Harold et al. 2009, Sleegers et al. 2010, Hollingworth et al. 2011, Naj et al. 2011, Lambert et al. 2013); in addition to the well documented APOE *ε4* allele, 19 GWAS loci were identified by Lambert et al. (2013) in the largest AD meta-analysis of GWAS to date (Table 1.1). Since then, polygenic risk scores (PRS) for AD have also been developed (Escott-Price et al. 2015). Collectively, common single nucleotide polymorphisms (SNPs) are believed to only account for 33% of the attributed risk of disease (Ridge et al. 2013) and the mechanisms behind their action remains largely unknown. Exome-sequencing projects have also identified other variants e.g. *TREM2* (Guerreiro et al. 2012), which have a larger effect size, yet are relatively rare in the standard population.

By far the most robust genetic loci associated with LOAD development risk is the APOE *ε4* allele. The rate of β-amyloid monomer aggregation is thought to be affected by *APOE* genotype status. Inheritance of two *APOE ε4* alleles infers a considerably higher risk of inheriting AD at some point in a person's lifetime (Corder et al. 1993). However the inheritance of the *APOE ε2* allele actually infers a protective effect from neurodegenerative disease (Corder et al. 1994) and *APOE ε3* has no effect on disease risk. Hypotheses as to the role the APOE protein plays in the process of β-amyloid congregation are varied and contradictory. The first theory revolves around APOE's role as the major protein component of very low density lipoproteins (vLDL), the major apolipoprotein within the brain. Research has shown that high levels of LDL cholesterol in the brain correlate with AD pathology post-mortem (Kuo et al. 1998). Specifically, the quantity of cholesterol in the neuronal plasma membrane has been shown to make neurones more susceptible to the damage caused by β-amyloid in AD (Arispe et al. 2002). Other studies suggest that

cholesterol acts directly on the amyloid cascade by promoting pathogenic processing of *APP* (Mattson 2004); work in *APP* mutant transgenic mice has shown that cholesterol metabolism can increase the production of β-amyloid (Puglielli et al. 2003). Another hypothesis regarding the role APOE plays in AD pathology revolves around its ability to clear amyloid deposits in certain conditions based on APOE's concentration within the brain/cerebrospinal fluid (CSF) (Cruchaga et al. 2012). The concentration of proteins is thought to vary as a function of the *APOE* allele(s) inherited. However, studies have shown that the presence of the *APOE ε4* allele leads to lower concentrations of the resulting protein in the plasma, but no significant change in CSF levels (Rezeli et al. 2015), with further work showing that plasma APOE levels have no effect on disease susceptibility (Simon et al. 2012). Recent work by Chung et al. (2016) supports an alternative theory, by showing that APOE isoforms have profound effects on the accumulation of complement protein C1q in ageing brain. C1q is an important factor in the rate of astrocytic turnover of synapses. More specifically it has been found that the *APOE ε2* allele increased synaptic phagocytosis by astrocytes, whereas *APOE ε4* caused a decrease, potentially leading to the build-up of senescent synapses, leading to neurodegenerative diseases such as AD (Chung et al. 2016).

SNP	Chr.	Position	Gene Name	Alleles	Meta P
rs6733839	2	127892810	*BIN1*	C/T	6.9×10^{-44}
rs10792832	11	85867875	*PICALM*	G/A	9.3×10^{-26}
rs9331896	8	27467686	*CLU*	T/C	2.8×10^{-25}
rs6656401	1	207692049	*CR1*	G/A	5.7×10^{-24}
rs983392	11	59923508	*MS4A6A*	A/G	6.1×10^{-16}
rs4147929	19	1063443	ABCA7	G/A	1.1×10^{-15}
rs11218343	11	121435587	*SORL1*	T/C	9.7×10^{-15}
rs28834970	8	27195121	*PTK2B*	T/C	7.4×10^{-14}
rs11771145	7	143110762	*EPHA1*	G/A	1.1×10^{-13}
rs9271192	6	32578530	*HLA-DRB5– HLA-DRB1*	A/C	2.9×10^{-12}
rs10948363	6	47487762	*CD2AP*	A/G	5.2×10^{-11}
rs1476679	7	100004446	*ZCWPW1*	T/C	5.6×10^{-10}
rs2718058	7	37841534	*NME8*	A/G	4.8×10^{-9}
rs10498633	14	92926952	*SLC24A4 / RIN3*	G/T	5.5×10^{-9}
rs17125944	14	53400629	*FERMT2*	T/C	7.9×10^{-9}
rs10838725	11	47557871	*CELF1*	T/C	1.1×10^{-8}
rs7274581	20	55018260	*CASS4*	T/C	2.5×10^{-8}
rs35349669	2	234068476	*INPP5D*	C/T	3.2×10^{-8}
rs190982	5	88223420	*MEF2C*	A/G	3.2×10^{-8}

Table 1.1; Genome wide significant genetic variants associated with AD. *The largest GWAS meta-analysis of LOAD identified 19 loci that reached genome wide significance ($P < 5x10^{-8}$) (Lambert et al. 2013).*

1.1.1.4. Non-genetic mechanisms

Back in 2001 the “Challenging Views of Alzheimer's Disease” meeting discussed the proposition that environmental agents, such as diet, aluminium, and viruses, are as important as genetic factors in the aetiology of AD. Dietary levels of fat were found to be a significant risk factor for the development of AD in a dozen countries (Grant et al. 2002). This could be linked to the amyloid cascade hypothesis where it is suggested that high levels of CSF cholesterol can cause neurones to be more susceptible to β-amyloid derived damage. Diet, aluminium, and viral infections may increase the prevalence of AD through action of the inflammatory hypothesis, eliciting inflammation, which may cause the neurological damage that result in AD. A prospective analysis of risk factors for AD was performed in 2002, and not surprisingly found increasing age and fewer years of education to be the main risk factors for AD (Lindsay et al. 2002). Interestingly, no statistically significant association was found for family history of dementia, sex, history of depression, oestrogen replacement therapy, head trauma, antiperspirant or antacid use, smoking, high blood pressure, heart disease, or stroke.

In recent years researchers have used epigenome-wide association studies (EWAS) to identify epigenetic changes in disease with the aim of elucidating additional novel mechanisms involved in aetiology, which may also provide a link between genetic and environmental factors. Epigenetic processes mediate the reversible regulation of gene expression, occurring independently of DNA sequence variation, acting principally through chemical modifications to DNA and nucleosomal histone proteins (Lunnon et al. 2013). Dynamic changes to the epigenome orchestrate a diverse range of important neurobiological and cognitive processes in the brain (Lunnon et al. 2013). Recent advances in genomic technology have allowed the first genome-scale studies assessing methylomic variation in AD. These studies have identified AD-associated DNA methylomic variation at numerous loci in the cortex, with consistent findings across multiple independent study cohorts, in addition to brain-region specific changes and blood DNA methylation signatures (De Jager et al. 2014, Lunnon et al. 2014), which are discussed in more detail in Section 1.2.5.

1.2. Epigenetics

"Epi" literally means above, therefore, "epigenetics" is essentially information that is coded above the level of the standard genetic code. As such, epigenetics refers to alterations in gene expression, which are brought about by heritable, but potentially reversible, changes in chromatin structure and/or DNA modifications (Steven Henikoff et al. 1997). Although every nucleated cell in the human body contains the same genetic information, no two cell types are physiologically the same. This is due to a complex interplay of epigenetic mechanisms that regulate the expression of genes on a cellular basis. This regulation can be altered by a number of different stimuli. As such, epigenetic mechanisms are dynamic, allowing the cell to adapt to its environment without changing its underlying genetic code (Jaenisch et al. 2003) . However, this change can also be the cause of disease with many studies showing that changes in the epigenetic landscape can lead to various forms of cancer (Jones et al. 1999, Feinberg et al. 2004, Laird 2005, Esteller 2008, Rauch et al. 2012), and such changes have been hypothesised to play a role in various neurological diseases, such as AD (Mill 2011). As epigenetic modifications are transient, it gives credence to the idea that if an epigenetic cause for disease can be identified, reversing this change could be a feasible treatment method.

1.2.1. Cytosine modifications

DNA methylation, which modulates the transcription of mammalian genomes, is the best characterised and most stable epigenetic modification. This is due to its ability to be interrogated using archived genomic DNA resources, which are the focus of most human epidemiological epigenetic research to date (Lunnon et al. 2013). The methylation of a cytosine in a CpG dinucleotide, forming 5-methylcytosine (5mC), can disrupt the cell's transcriptional machinery by blocking the binding of transcription factors and attracting methyl-binding proteins that initiate chromatin compaction and bring about gene silencing (Klose et al. 2006). The predominant focus to date is methylation within CpG Islands (CGIs), dense regions of CpG sites, located within the 5' promoters of many constitutively expressed housekeeping control genes. However, recent data suggests that the relationship between 5mC and transcription may be more complex, with gene body methylation and non-CpG

methylation often being associated with active gene expression (Hellman et al. 2007, Ball et al. 2009, Lister et al. 2009, Rauch et al. 2009, Aran et al. 2011) and alternative splicing (Lyko et al. 2010, Flores et al. 2012).

Further research has now demonstrated that other cytosine modifications not only exist, but may also have functional roles in gene regulation (Inoue et al. 2011, Ito et al. 2011). These modifications include 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), which occur as stages in the de-methylation of cytosine (Breiling et al. 2015) (Figure 1.5). DNA Methyl transferase (DNMT) catalyses the addition of a methyl group to the fifth carbon of cytosine to generate 5mC. During demethylation 5mC is oxidised to 5hmC by the Ten-eleven translocation (TET) family of enzymes (Tahiliani et al. 2009, Ito et al. 2010). TET enzymes continue to oxidise 5hmC to 5fC and on to 5caC (Ito et al. 2011). 5fC and 5caC can then be converted to unmodified cytosine following base-pair excision / Terminal deoxynucleotidyl transferase (TdT) (Guo et al. 2011). As such, these modifications were previously considered to only be transient changes, having no effect on gene transcription/regulation. It has been suggested that DNA replication leads to a passive removal of 5hmC from the genome (Inoue et al. 2011), however work by Hashimoto et al. (2012) has shown that the DNA methyl transferase 3B (DNMT3B) shows an equal affinity for un-methylated, hemi-methylated or hemi-hydroxymethylated DNA post replication, leading to the theory that 5hmC is also maintained during cellular replication (Hashimoto et al. 2012). Although considerably less abundant in the genome (Breiling et al. 2015), these modifications, especially 5hmC, have recently been shown to be more stable than previously thought (Bachman et al. 2014). Recent research has found that 5hmC in particular has a significantly higher abundance in the brain compared to other tissues (Lister et al. 2013). 5hmC was also found to be elevated and highly dynamic during foetal development (Wang et al. 2012, Spiers et al. 2017). These studies, coupled with advances in research methodologies, mean that 5hmC, and other modifications, are starting to be researched in the context of complex diseases, with recent studies demonstrating that 5hmC and 5mC have opposing effects on gene expression (Sherwani et al. 2015).

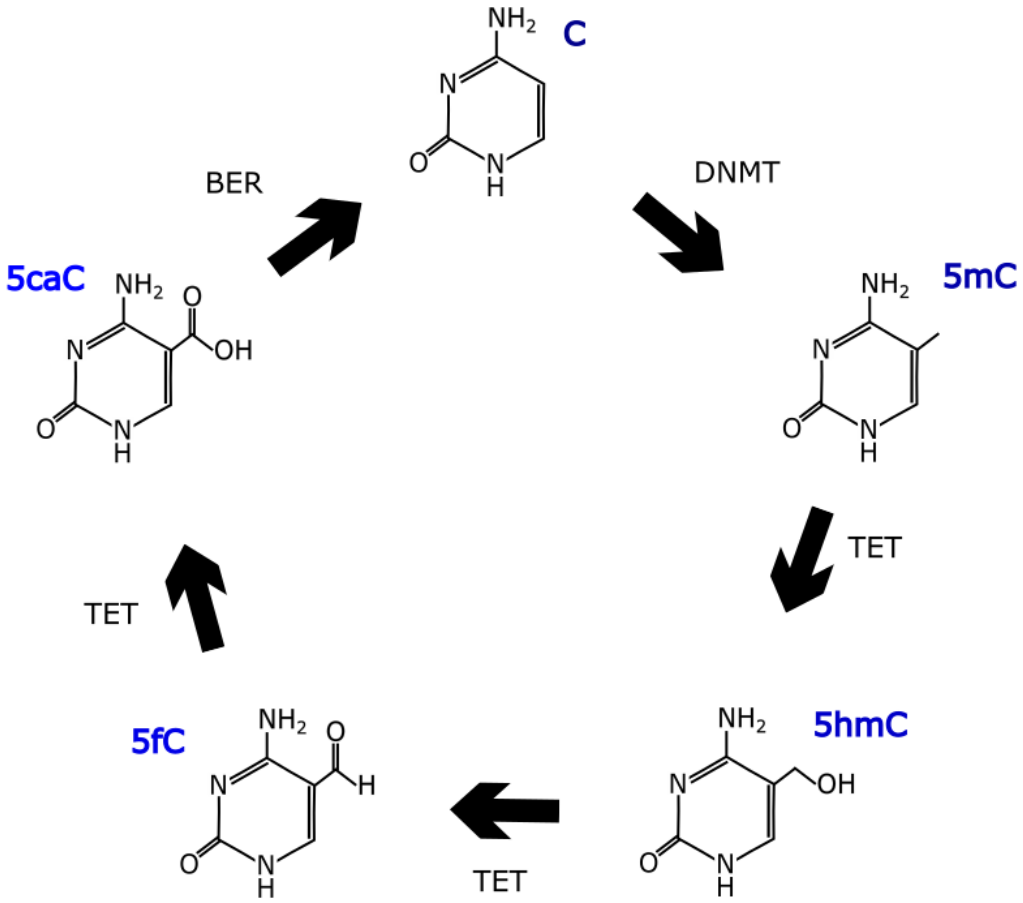


Figure 1.5; Cytosine modifications as stages of the de-methylation process. *Cytosine can be methylated to 5mC by DNA methyltransferase (DNMT). 5mC can then be oxidised to 5hmC, 5fC and 5caC by ten-eleven translocation (TET) dioxygenases and through a process of base excision repair (BER) is returned to cytosine.*

1.2.2. Histone modifications

Cytosine modifications are by no means the full extent of the epigenetic mechanisms that regulate gene expression. A complex interplay of DNA wound around octamers of core histone proteins to form nucleosomes, the basic component of chromatin (Figure 1.6), also plays a key role. This complex three-dimensional structure is found in the nuclei of all eukaryotic cells. The wrapping of the double stranded DNA helix around histone proteins enables the 10,000-20,000 fold level of compaction necessary to fit a genome into the small volume of the nucleus (Zentner et al. 2013). The level of chromatin compaction and therefore transcriptional ability is regulated by conformational changes to histone proteins including modifications such as methylation, phosphorylation, acetylation, ubiquitination, sumoylation, citrullination and adenosine diphosphate-ribosylation as well as other post-translational modifications of the amino acids that make up histone protein structure (Sadakierska-Chudy et al. 2015). These histone "tail" modifications regulate the level of accessibility of transcription start sites (TSS) and promotor regions within the DNA to transcription factors and other necessary cellular machinery required for transcription. This is done by shifting the chromatin from a condensed state (heterochromatin) to a more relaxed form (euchromatin). Acetylation and methylation are the most widely studied histone modifications (Table 1.2). Histone acetylation appears to be a ubiquitous marker of gene expression (Kim 2014). The second most studied histone modification is methylation. The methylation of histones has differing effects on gene transcription based on which amino acid within the histone is methylated and the number of methylation groups per amino acid. Further details on histone methylation can be found in Section 5.1.

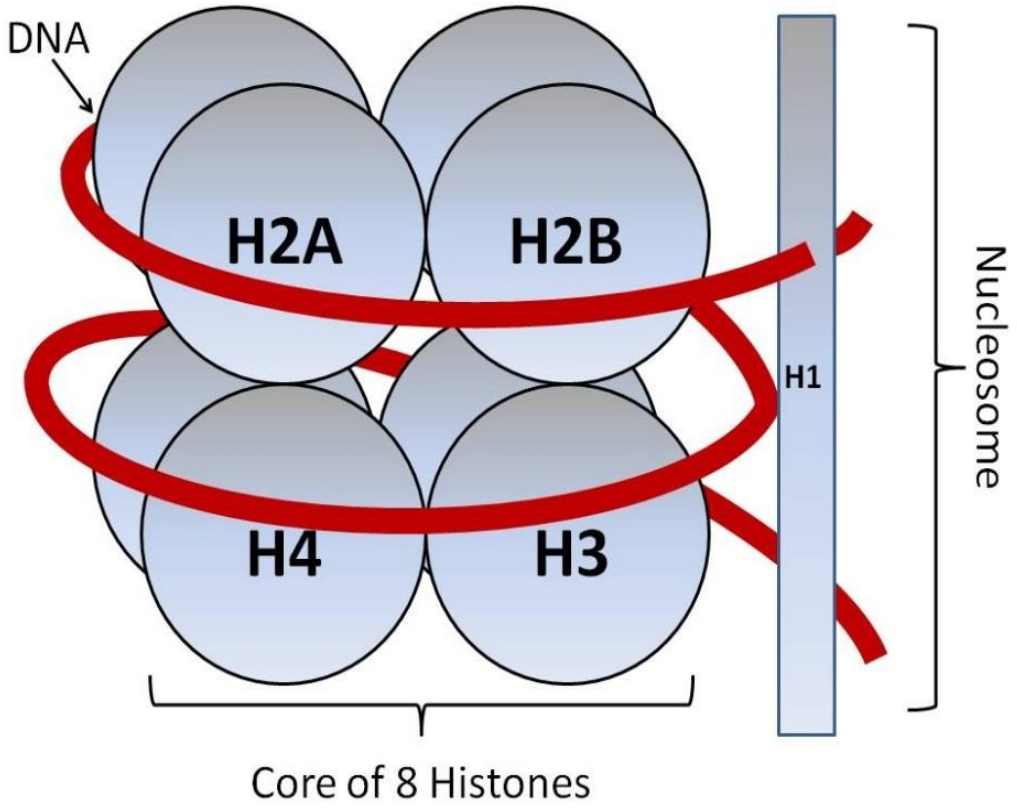


***Figure 1.6; Nucleosome, the core component of chromatin**. DNA is wrapped around the octameric nucleosome core of histone proteins which exist as dimers (H2A & H2B, H3 & H4) and then bound in position by a fifth linker histone protein monomer (H1/H5). Image taken from (Popa 2017)*

	Modification			
Position	**Mono-Methylation**	**Di-Methylation**	**Tri-Methylation**	**Acetylation**
H2AK5	-	-	-	Activation
H2AK7	-	-	-	Activation
H2AK9	-	-	-	Activation
H2AK13	-	-	-	Activation
H2BK5	-	-	-	Activation
H2BK15	-	-	-	Activation
H2BK20	-	-	-	Activation
H2BK120	-	-	-	Activation
H3R2	Activation	-	-	-
H3K4	Activation	Activation	Activation	-
H3K9	Activation/ repression	Repression	Activation/ repression	-
H3K14	-	-	-	Activation
H3R17	Activation	-	-	-
H3K18	-	-	-	Activation
H3K23	-	-	-	Activation
H3R26	Activation	-	-	-
H3K27	Activation	Repression	Repression	Activation
H3K36	-	Activation	Activation	-
H3K56	-	-	-	Activation
H3K79	Activation	Repression	Repression	
H3K115	-	-	-	Activation
H4R3	-	Activation	-	-
H4K5	-	-	-	Activation
H4K8	-	-	-	Activation
H4K12	-	-	-	Activation
H4K16	-	-	-	Activation
H4K20	Activation/ repression	Repression	Repression	Activation
H4K59	Repression	-	-	-
H4K91	-	-	-	Activation

Table 1.2; Major modifications of histone proteins and their impact on genetic function. *Taken from (Kim 2014)*

Histone modifications have been shown to be largely restricted to transcribed regions of the genome, suggesting that their function is tightly linked to polymerase activity (Schübeler et al. 2004). Histone modifications occur in a highly cell-type-specific manner at enhancers; interestingly however, there appears to be similar patterns of histone modifications at promotors between cell types. This therefore suggests that enhancers are the most variable class of transcriptional regulatory element and probably drives the cell-type-specific gene expression patterns (Heintzman et al. 2009). In heterochromatin, low levels of acetylation, methylation and phosphorylation can be detected on genes, however these levels are not sufficient to cause transcription. Typically euchromatin has higher levels of acetylation and is tri-methylated at H3K4, H3K36 and H3K79. The methylation of histone sites is typically associated with changes in gene transcription and DNA repair (Kouzarides 2007).

1.2.3. Micro-RNAs

Another interesting field of epigenetic regulation is the study of micro-RNAs (miRNAs), which also play a critical role in gene regulation. Currently 2588 mature miRNAs are listed in release 21 of miRBase in the human genome across multiple human tissues, including the brain (Kozomara et al. 2014). The small 19-26 nucleotide RNAs often originate from their own genes with their own promoter and regulatory regions (Bartel 2004). miRNAs regulate gene expression post-transcriptionally by binding to their target mRNAs and inhibiting translation or, less often, inducing cleavage of the mRNAs (Bartel 2004).

1.3. Epigenetic studies in AD

Although many of the genomic studies to date have identified robust and reproducible findings of genetic variation linked to AD pathology, these do not account for all of disease incidence. This, coupled with a number of disease attributes, are suggestive of an epigenetic contribution to pathology. For example, differential vulnerability of specific brain regions to pathology, the age of disease onset, environmental influences on disease and the absence of a statistically significant association of family history of dementia (Lindsay et al. 2002, Profenno et al. 2010). There have been numerous studies into the epigenetics of AD, for a detailed review see Smith et al. (2017). To date, three EWAS of DNA methylomic variation, in AD transgenic mice, have been published (Sanchez-Mut et al. 2013, Agbemenyah et al. 2014, Cong et al. 2014). Sanchez-Mut et al. (2013) identified a number of genes that were differentially methylated in diseased mouse brain, including *Tbxa2r*, *F2rl2*, *Sorbs3* and *Spnb*. They were also able to translate these findings to patients with AD, identifying 5mC-associated silencing of *TBXA2R*, *SORBS3* and *SPTBN*. Using methylated DNA immunoprecipitation (MeDIP)-chip Cong et al. (2014) and Agbemenyah et al. (2014) identified over 2000 hypermethylated CpG sites. However, there are limitations to the utility of rodents for modelling human AD, for example mice do not naturally get AD symptoms or produce amyloid plaques; therefore transgenic mouse models of amyloid pathology are due to them producing human amyloid. Currently no mouse model has all the features of human LOAD and these models are currently based on genetic mutations associated with familial AD. Therefore these mice simply model the effects of the accumulation of β-amyloid plaques and/or NFTs, rather than being a model of sporadic AD, which accounts for ~95% of cases. However, the majority of epigenetic studies of AD have focussed on studying the human disease in post-mortem brain samples. As this is a new research field, to date epigenetic studies in AD post-mortem brain have been generated by only a handful of researchers in a small number of research teams, this source of potential bias will resolve as research interest grows.

1.3.1. DNA Methylation in AD

Attempts to characterise the global DNA methylome in AD brain using antibodies has shown reductions in 5mC across numerous brain regions (Mastroeni et al. 2009, Mastroeni et al. 2010, Chouliaras et al. 2013). However, other studies have not supported these findings (Bradley-Whitman et al. 2013, Condliffe et al. 2014, Coppieters et al. 2014, Lashley et al. 2015). Candidate gene studies aiming to identify AD-associated 5mC changes have been carried out in human blood (Bollati et al. 2011, Arosio et al. 2012, Furuya et al. 2012, Furuya et al. 2012, Di Francesco et al. 2013, Hernandez et al. 2014) and brain tissue (Sun-Chong Wang 2008, Barrachina et al. 2009, Brohede et al. 2010, Iwata et al. 2014, Silva et al. 2014). These studies targeted many of the genes nominated from genetic studies of familial AD, in addition to enzymes involved in DNA methylation, for example DNMT1.

More recently, advances in epigenomic technology have allowed the quantification of DNA methylomic variation across the genome in AD. The first EWAS of AD (Bakulski et al. 2012) used the Illumina Infinium 27K Beadarray to quantify 5mC at ~27,000 CpG sites in the frontal cortex of 12 late-onset AD donors and 12 cognitively normal matched control subjects. Despite low sample numbers their most significant loci was found in the *TMEM59* gene, which is thought to be involved in *APP* processing. Sanchez-Mut et al. (2014) also used the 27K Beadarray to examine hippocampal samples from five control donors to those with early-stage AD (Braak I-II), mid-stage AD (Braak III-IV) and late-stage AD (Braak V-VI), demonstrating a >25% methylation difference between controls and Braak stage V–VI at single CpG sites within three genes: *CLDN15, QSCN6* and *DUSP22*, however these results have not been replicated since. Since then, the development of the Illumina Infinium 450K Beadarray (450K array), has allowed researchers to profile ~485,000 CpG sites. De Jager et al. (2014) used this technology in a large cohort of 708 prefrontal cortex samples to examine DNA methylomic differences associated with neuritic plaque burden. They identified 71 differentially methylated probes, 11 of which were replicated by Lunnon et al. (2014). Furthermore, altered gene expression with AD pathology was found in seven of the replicated differentially methylated genes in separate temporal cortical samples, including the *ANK1* gene. Simultaneously, Lunnon et al. (2014) performed a cross-tissue analysis of 5mC changes associated

with Braak Stage using the 450K Beadarray in a discovery cohort of 117 individuals. Combining EWAS data from different cortical brain regions from the same donors, they were able to identify a number of consistently differentially methylated loci, which included CpG sites in the *ANK1* gene. They confirmed this region of Braak-associated DNA hypermethylation in *ANK1* in an independent 450K EWAS dataset of 146 cortical brain samples, and in a separate replication cohort of 62 samples, where they used pyrosequencing to look at an extended region of eight CpG sites in the *ANK1* gene. This showed that the differentially methylated region (DMR) in *ANK1* in AD cortex spanned at least six CpG sites. Watson et al. (2016) also used the 450K Beadarray platform to identify AD-related 5mC changes in the superior temporal gyrus in 34 AD cases and 34 non-demented controls matched by age at death, race, and sex. They also showed overlap between their most significant DMRs and those loci nominated by the Lunnon *et al.* and De Jager *et al.* studies. Since, a number of further studies have built on these now publically available EWAS datasets, looking for associations in genetic risk variants for AD from GWAS (Chibnik et al. 2015), or exome sequencing studies (Smith et al. 2016).

1.3.2. Other epigenetic modifications in AD

To date studies of 5hmC in AD brain have been largely limited to global profiling methods, with conflicting results. Two studies have shown global decreases in 5hmC in the hippocampus (Chouliaras et al. 2013), EC and CER (Condliffe et al. 2014), with another study showing increased 5hmC in the middle frontal gyrus and middle temporal gyrus (Coppieters et al. 2014), whilst a fourth study revealed no difference in the EC (Lashley et al. 2015). Previous publications, described in Section 1.3.1, that used bisulfite treatment to assess DNA methylation have in fact, profiled the combination of both 5mC and 5hmC as a combined measure. This is due to 5hmC reacting with bisulfite to form 5-methylenesulfonate which, similarly to 5mC, does not undergo C-to-T transition following PCR (Huang et al. 2010). This means previous 5mC results are confounded by 5hmC suggesting that genes of interest have been missed. One recent study claimed to have undertaken the first EWAS of 5hmC in AD used a selective chemical labelling technique to enrich for 5hmC, and then sequenced captured libraries from a small cohort of 30 individuals with either AD, mild cognitive impairment (MCI) or no dementia (Zhao et al. 2017). However, due to

low sequencing resolution, this study's analysis was unable to discriminate between 5hmC and 5mC and as such; an EWAS of 5hmC in AD is still required.

Histone modifications have also been studied in respect to AD pathology. Two studies, looking at APP/PS1 transgenic mice and human post-mortem temporal lobe respectively, have identified a significantly lower amount of histone acetylation at a global level (Francis et al. 2009, Zhang et al. 2012). However, work by Walker et al. (2013) has found higher H3 and H4 acetylation levels in neurones from a 3x transgenic AD mouse model (*PS1*, *APP* and *MAPT*) compared to controls. Finally, Rao et al. (2012) found global histone phosphorylation in post-mortem frontal cortex, although they only analysed a small (N = 20) number of individuals. In summary, although there has been a rapid increase in the amount of epigenetic research in AD, further research is required to establish the role of 5hmC and histone modifications in the disease.

1.4. Integration of genetic and epigenetic studies to identify dysfunctional AD pathways

Although GWAS and EWAS analyses have identified multiple genes associated with AD, with the exception of *BIN1* the genetic location of loci identified do not overlap. However, a number of overlapping pathways have been identified between published GWAS and EWAS studies, highlighting dysfunctional synaptic, lipid metabolism, plasma membrane/cytoskeleton, mitochondrial and immune cell activation pathways (Smith et al. 2016). The identification of common pathways altered in genetic and epigenetic studies will aid our understanding of disease mechanisms and identify potential novel targets for pharmacological intervention.

1.4.1. Plasma membrane/cytoskeleton

We have previously reported that the plasma membrane/cytoskeletal pathway contains the largest number of loci nominated from published GWAS and EWAS studies (Smith et al. 2016). The plasma membrane insulates the intracellular components from the extracellular environment, as well as catalysing the transport of specific compounds, including nutrients and ions. Phospholipids that make up the membrane provide suitable fluidity and permeability. Alterations in the receptor function, membrane integrity, and membrane-dependent processes seen in AD have been previously reviewed (Farooqui et al. 1995). The cytoskeleton provides contractility and couples biochemical responses with mechanical stresses in cells. It is vital in the movement of cellular machinery around the cell and to the membrane, as well as orchestrating the procedures needed for cellular movement and re-shaping, a function important in the microglial cells of the brain in the response to inflammatory stimuli (Sheng et al. 1997). For an overview of cell mechanics and the cytoskeleton see the review by Fletcher et al. (2010). The inability of neurones to regulate calcium homeostasis through cell surface ion channels is an aspect of AD pathogenesis that appears to be intimately involved in the dysfunction and death of neurones (Mattson 2004). Familial AD mutations in *APP* and *PSEN1* support a role for perturbed calcium regulation in AD (Mattson 2004). In addition, all of the enzymatic machinery responsible for the generation of the pathogenic β-amyloid

species are plasma membrane based (Lukiw 2013); suggesting that damage to the plasma membrane may be the cause of β-amyloid pathology typical of AD.

BIN1 has been nominated by both GWAS and EWAS, and in addition to its role in synaptic signalling, it also has a role in plasma membrane/cytoskeletal processes as it acts as an amphiphysin, which are known to promote caspase-independent apoptosis as well as play an important role in neuronal membrane organisation (Wigge et al. 1997). Major learning defects and seizures have been linked to decreased expression of amphiphysins in murine brain (Di Paolo et al. 2002). In addition, altered expression of *BIN1* has been shown in ageing mouse models of AD (Yang et al. 2008), providing further evidence for its role in AD pathology. Despite having no previous link to AD, *ANK1* is now one of the strongest reported candidate genes in AD EWAS as described in Section 1.3.1, with strong links to cell structure. *ANK1* is found in multiple different isoforms and is expressed at different levels throughout the brain (Gallagher et al. 1997), with some evidence for differential splicing in AD (Lunnon et al. 2014). As with *BIN1,* one of the main functions of *ANK1* is compartmentalisation and maintenance of the plasma membrane and it is possible that the altered expression of this gene could lead to neuronal membrane dysfunction in AD (Lunnon et al. 2014).

The *PVRL2* gene identified by GWAS encodes a single-pass type I membrane glycoprotein, which is one of the plasma membrane components of adherens junctions. Cell to cell connections brought about by adherens junctions are vital for effective neuronal signalling (Marambaud et al. 2002). Interestingly, Marambaud et al. (2002) used various immunological based methods to investigate the *PSEN1*/γ-secretase system, where mutations are associated with familial AD, and showed it disrupted adherens junctions in AD. Expression of *PVRL2* has been detected in many organs including the brain, and it was later suggested it was associated with human longevity along with the AD GWAS nominated loci *TOMM40* and *APOE* (Lu et al. 2014). In addition, Elias-Sonnenschein *et al.*, showed a significant correlation between the GWAS nominated locus *MS4A4A* and β-amyloid pathology, but not with tau pathology in AD (Elias-Sonnenschein et al. 2013). Despite this, there is little-to-no research on the specific function of *MS4A4A*, although the gene product is associated with GO pathways that indicate it is an integral component of the plasma

membrane. However, GO pathways have limited value as they are based on incomplete gene annotations, with a large number of genes that are not linked to any pathway (Khatri et al. 2012).

1.4.2. Lipid homeostasis

Recent epidemiological, molecular and biochemical evidence has strengthened the hypothesis that cholesterol is a risk factor for AD, and although cholesterol homeostasis in the brain is largely unexplored, new findings strongly support the involvement of cholesterol in both the generation and deposition of β-amyloid (Puglielli et al. 2003). Specifically, the quantity of cholesterol in the neuronal plasma membrane has been shown to make neurones more susceptible to the damage caused by β-amyloid in AD (Arispe et al. 2002). Other studies suggest that cholesterol acts directly on the amyloid cascade by promoting amyloidogenic processing of *APP* (Mattson 2004). Interestingly, statins, which are a class of cholesterol-lowering drugs, decrease β-amyloid levels as well as plaque deposition in *APP* transgenic mouse models (Fassbender et al. 2001). In addition, high cholesterol levels and changes to cholesterol metabolism can increase the production of β-amyloid in cell culture and mouse models (Puglielli et al. 2003). Three of the most significant genes identified in AD GWAS are associated with lipid metabolism (*APOE, APOC1, CLU*). *APOE* was first identified as a genetic risk factor for AD in 1993 (Strittmatter et al. 1993), using immuno-staining and genotyping analysis of 30 AD cases and 91 controls. Since 2006, and the wide application of GWAS to AD research (Grupe et al. 2007), the *APOE* polymorphism has been successfully replicated in several other studies (Coon.; et al. 2007, Abraham et al. 2008, Kramer et al. 2011, Logue et al. 2011, Meda et al. 2012, Ramanan et al. 2014), making *APOE* the most robust gene linked to LOAD risk to date. The proportion of genetic variance for LOAD risk attributed to *APOE* genotype is estimated to be 10–20% (Slooter et al. 1998). *APOE* is a 299 amino acid glycoprotein and the major protein component of vLDL, the major apolipoprotein in the brain (Puglielli et al. 2003), as well as having a functional role in cholesterol and triglyceride metabolism (Breslow et al. 1982). There are three *APOE* alleles that affect one's risk of developing AD (ε2, ε3 and ε4), in addition to age of onset (Roses 1996). Of the three alleles; *APOE* ε2 demonstrates a protective effect, with an OR of

0.3 for possessing one ε2 allele, whilst *APOE* ε4 is associated with a higher LOAD risk, with an OR of 4.4 and 19.3 respectively for having one or two alleles (Corder et al. 1994), as well as a younger median age of dementia onset (Corder et al. 1994, Reiman et al. 2007). It has been suggested that the mutated *APOE* hinders clearance of soluble β-amyloid protein from the brain, leading to the hallmark β-amyloid aggregation into fibrils. Furthermore, *APOE* has been shown to promote neurodegeneration by directing the toxic β-amyloid oligomers to synapses (Ramanan et al. 2014). However, a recent positron-emission tomography (PET) study to measure β-amyloid in 602 individuals found that the ε4 allele is neither necessary, nor sufficient, for the development of AD pathology (Ramanan et al. 2014).

SORL1 has been identified in several genomic studies of AD, using GWAS and exome sequencing methods (Rogaeva et al. 2007, Miyashita et al. 2013). In addition, Yu *et al*. found epigenetic changes in this gene when they investigated 5mC changes in candidate genes in AD EWAS data (Yu et al. 2014). It has many functional domains with different functions, including cargo transport, chaperone-like activity, signalling, and intracellular sorting (Jacobsen et al. 2001). When acting as a sorting receptor, the *SORL1* gene product protects *APP* from being directed to the endosome where it would be cleaved by β-secretase, producing β-amyloid (Louwersheimer et al. 2015). Further, SORL1 can bind APOE, making *SORL1* an important component in the pathophysiology of AD (Elias-Sonnenschein et al. 2013).

1.4.3. Synaptic signalling

Synaptic dysfunction is possibly the best-established of all the proposed pathological mechanisms for AD to date as it shows clear progression throughout the entire disease, including pre-symptomatic changes (Masliah et al. 2001). Early stages of AD are characterised by a 25-35% decrease in numerical density of synapses per cortical region (Davies et al. 1987). There has also been evidence that the loss of synapses correlates with the soluble pool of cortical β-amyloid (Lue et al. 1999). Stereological and biochemical analyses have shown that the reduction in synaptic density within the AD brain correlates with cognitive defects better than the traditional hallmarks of β-amyloid plaques and neurofibrillary tangles (Masliah et al. 2001).

Four genes from GWAS and EWAS analyses of AD are linked to synaptic function (Smith et al. 2016). Two of these, *BIN1* and *PICALM,* have functions in vesicular trafficking. Specifically, studies have shown that the *BIN1* gene has roles in a number of specific pathways, including clathrin-mediated endocytosis (CME) which is an essential step in the intracellular trafficking of proteins and lipids such as nutrients, growth factors and neurotransmitters in synapses (Dreyling et al. 1996, Wigge et al. 1998, Cousin et al. 2001). Originally identified as a tumour suppressor (Sakamuro et al. 1996), the *BIN1* gene product is expressed most abundantly in brain and muscle (Wechsler-Reya et al. 1997), with several alternatively spliced brain specific isoforms. *BIN1* is one of the few genes that has been reproducibly identified by GWAS that does not fall near or within the *APOE* locus, in addition it is the only gene to be significantly associated with AD in both GWAS and EWAS.

Like *BIN1*, *PICALM* is also involved in CME (Dreyling et al. 1996). PICALM directs the trafficking of the VAMP2 protein. VAMP2 is a SNARE protein that plays a key role in the fusion of vesicles to the presynaptic membrane allowing neurotransmitter release into the synapse, a process essential to neuronal function (Harel et al. 2008). *PICALM* has been robustly identified as a risk factor for AD via GWAS (Harold et al. 2009, Seshadri et al. 2010), however, AD linked SNPs identified in *PICALM* may still be affected by *APOE* genotype, due to the large amount of attenuation seen when adjusted for *APOE* status (Hu et al. 2011). Jun *et al.* have also reported this interaction observing that genotypes of *PICALM* conferred risk predominantly in *APOE* ε4–positive participants, providing strong evidence for a synergistic effect (Jun et al. 2010). PICALM is also thought to affect APP processing via endocytic pathways (Harold et al. 2009).

As a previously known risk factor gene for AD (Lambert et al. 2013), *PTK2B* was shown via network analysis to be linked to the recently identified loci in *RHBDF2*, *ANK1* and *RPL13* from AD EWAS, providing further evidence for a role in AD pathology (De Jager et al. 2014). *PTK2B* has a number of roles including the induction of long term potentiation of nerve cells, a central process of memory formation; cell migration and synaptic function (Lambert et al. 2013).

1.4.4. Immune cell dysfunction (astrocytes, oligodendrocytes and microglia)

There is a widely accepted link between inflammation, the immune system and AD pathology (Haruhiko Akiyama 2000, Tuppo et al. 2005, Salminen et al. 2009, Adriana Martorana et al. 2012, Monson et al. 2014) , more specifically the inflammation seen in AD has been proposed to exacerbate symptoms (Haruhiko Akiyama 2000). Microglia, which are the brain's resident macrophages, have been shown to increase their viability by 22.0~29.4% in response to fibrillar β-amyloid deposits of 0.2 to 5.0μM, which are commonly seen in AD. Oligomeric β-amyloid at a dose of 5.0μM results in cytotoxic microglia (Pan et al. 2011) and ultimately leads to synaptic degeneration and neuronal death (Barger et al. 2001). However, relatively few genes that have shown robust associations with AD have been directly linked with inflammation or immune functions. Most noteworthy a rare variant in *TREM2,* considerably enriched in AD, was recently recognised by a number of AD exome sequencing and GWAS projects (Forabosco et al. 2013, Guerreiro et al. 2013, Harald Neumann 2013, Jonsson et al. 2013). *TREM2* encodes an innate immune system receptor on the surface of microglial cells within the brain. With the signalling counterpart DAP12 (also called TYROBP) TREM2 forms a molecular complex that promotes phagocytosis of bacteria (N'Diaye et al. 2009). Work by Takahashi *et al*. has shown that TREM2 also has a role in the clearance of apoptotic neurones, due to its ability to increase migration and phagocytosis of microglia (Takahashi et al. 2005). Recently one study demonstrated correlation in *TREM2* and *CD33* gene expression in AD (Chan et al. 2015). As *CD33* has also been nominated in various AD GWAS (Hollingworth et al. 2011, Naj et al. 2011, Kamboh et al. 2012) this provides further evidence for an overlap of AD gene pathways in disease. As described above, recent protein-protein interaction data also demonstrated that several EWAS nominated loci (*ANK1, RHBDF2, PICLAM*) have a functional link to *PTK2B* (De Jager et al. 2014). *PTK2B* is an AD risk factor gene that plays a key role in the signalling cascade involved in the modulation of microglial and infiltrating macrophage cell activation (De Jager et al. 2014).

A further gene related to immune function is *RHBDF2*, identified by EWAS. Differentially methylated CpG sites close to the *RHBDF2* gene were identified in two independent EWAS (De Jager et al. 2014, Lunnon et al. 2014), which increases

RHBDF2 expression in AD brain (De Jager et al. 2014). *RHBDF2* transports TNFα converting enzyme (TACE, also called ADAM17), which is necessary for the release of TNFα from the cell surface (Adrain et al. 2012). *RHBDF2* absence in mice affects the release of TNFα from the cell surface (Siggs et al. 2012) and therefore impairs systemic immune responses to pathogens (McIlwain et al. 2012), although the brain phenotype has yet to be researched.

1.4.5. Mitochondrial processes

Mitochondrial dysfunction is one of the most prominent characteristics of AD, in both the brain and the periphery (Lunnon et al. 2012, Devall et al. 2014, Devall et al. 2016), with one of the most robust genes identified from GWAS associated with mitochondrial function (*TOMM40*). This gene is located approximately 2kb downstream from *APOE* and due to the locality of these two genes there is strong linkage disequilibrium (LD) for *TOMM40* with the *APOE* locus (Feulner et al. 2010), hence many studies have failed to find an association of *TOMM40* in AD after adjusting for *APOE* genotype (Yu et al. 2007, Wijsman et al. 2011, Ramanan et al. 2014). However, one study reports *TOMM40* as a possible risk factor for AD independent of *APOE* genotype (Lutz et al. 2010). Specifically this study found a poly-T track mutation in *TOMM40* that acts independently of *APOE* genotype, which has also seen been reported in another independent study (Cruchaga et al. 2011). In addition to increasing risk of developing AD, *TOMM40* has also been linked to an earlier age of onset for the disease (Roses et al. 2009). Other studies also suggest that *TOMM40* provides an additional risk for AD, in addition to *APOE* (Potkin et al. 2009, Takei et al. 2009). However, until the extent of the LD between *TOMM40* and *APOE* is fully characterised, it will be difficult to pinpoint the exact effect the *TOMM40* mutation has on LOAD pathogenesis.

CLU has various nuclear and mitochondrial isoforms and is thought to regulate the rate of cell proliferation. *CLU* has been consistently replicated across many GWAS and holds strong association with AD (Harold et al. 2009, Lambert et al. 2009, Jun et al. 2010, Wijsman et al. 2011). The nuclear isoforms result in the promotion of apoptosis, whereas mitochondrial isoforms of *CLU* suppress BAX-dependent release of cytochrome c into the cytoplasm and inhibit apoptosis (Zhang et al. 2005). As an

increased level of apoptosis in the brain is seen in AD, it could suggest a role of *CLU* mutations in pathogenesis (Behl 2000). *SPG7* was identified by EWAS and encodes a mitochondrial metalloprotease protein. Mitochondrial proteases degrade misfolded and non-assembled polypeptides. They also regulate the activity of specific substrates by mediating essential processing steps. These proteases have been hypothesised to play a role in neurodegenerative diseases by affecting neuronal maintenance and axonal function (Martinelli et al. 2010).

1.5. The *ANK1* gene

As previously described, some of the most robust AD-associated differentially methylated probes (DMPs) reside within the *ANK1* gene (Smith et al. 2016). Lunnon *et al*. identified two CpG loci, residing 90bp apart, within exon 42 of the *ANK1* gene (longest isoform), which were hypermethylated in multiple cortical brain regions from AD donors in their EWAS. Using pyrosequencing, they demonstrated the DMR extended across 120bp and included at least six CpG sites (Lunnon et al. 2014). Interestingly this pattern of hypermethylation was not seen in the CER, a region typically devoid of AD pathology, even in the latest stages of the disease. In addition Lunnon *et al.* looked at the levels of expression of different *ANK1* transcripts and found expression levels of specific isoforms to be correlated with neuropathology.

The *ANK1* gene, located on the short arm of chromosome 8, belongs to the ankyrin family which was first discovered in erythrocytes, but since has also been found to be expressed in muscles and in the brain. *ANK1* is believed to link the integral membrane proteins to the underlying spectrin-actin cytoskeleton and play key roles in activities such as cell motility, activation, proliferation, contact, and maintenance of specialised membrane domains. Multiple isoforms of ankyrin with different affinities for various target proteins are expressed in a tissue-specific, developmentally regulated manner. Most ankyrins are typically composed of three structural domains: an amino-terminal domain containing multiple ankyrin repeats; a central region with a highly conserved spectrin-binding domain; and a carboxy-terminal regulatory domain, which is the least conserved and subject to variation. The small *ANK1* (*sAnk1*) protein splice variants make contact with obscurin, a giant protein surrounding the contractile apparatus in striated muscle. However, the precise function of most of the *ANK1* isoforms is not known. Alternative polyadenylation accounting for the different sized erythrocytic *ANK1* mRNAs has also been reported (Gallagher et al. 1997), in addition, muscle-specific isoforms of *ANK1* resulting from usage of an alternate promoter have also been identified (Gallagher et al. 1998). 50% of patients with the blood based disorder Hereditary Spherocytosis (HS) have a mutation in the *ANK1* gene (Randon et al. 1997, Miraglia del Giudice et al. 1998, Hughes et al. 2011). The HS disease results in the erythrocytes of the patient losing their biconcave shape and becoming completely

spherical. As a result the maximum oxygen carrying ability of the erythrocytes is reduced. In addition SNPs in *ANK1* have been shown to be associated with Type 2 Diabetes Mellitus (Imamura et al. 2012, Harder et al. 2013), a disease that is considered by many as a risk factor for AD (Luchsinger et al. 2001, Arvanitakis et al. 2004, Janson et al. 2004, Li et al. 2007, Profenno et al. 2010). Interestingly, in a population of Han-Chinese, a SNP within *ANK1* was found to be associated with AD (Chi et al. 2015). Most recently however, work by Mastroeni et al. (2017) has shown a four-fold increase in the expression of *ANK1* in laser-capture dissected microglia from AD brain compared with control. This highlights the importance of analysing epigenomic and transcriptomic changes in specific cell types in complex neurological disorders like AD.

1.6. Conclusions

To conclude, it is clear that a complex interplay of different factors, including genetic, environmental and epigenetic mechanisms, play a role in the aetiology of AD. The use of GWAS studies to identify common disease variants in AD has been at the forefront of research to understanding disease aetiology for ten years but a large proportion of disease prevalence is yet to be explained. More recently, the falling cost of exome and whole genome sequencing has identified rarer variants with a larger effect size. In contrast to the large amount of genetic data available for AD, epigenetic research in AD is still in its early stages, with just five EWAS of DNA methylomic variation in AD having been reported to date (Bakulski et al. 2012, De Jager et al. 2014, Lunnon et al. 2014, Sanchez-Mut et al. 2014, Watson et al. 2016). However, all of these studies have used bisulfite (BS)-treated DNA, which means that the data generated actually represents a sum of two different cytosine modifications (5mC and 5hmC). As such, estimates of 5mC in these studies may have been confounded by 5hmC levels and so may be an under or over-representation of actual levels of 5mC in AD at specific loci. Further work is now needed to profile 5mC and 5hmC independently in AD and to further characterise epigenetic changes in *ANK1* in AD brain, for example by examining the specificity of the DMR across different dementias and by investigating different DNA and histone modifications across the gene.

1.7. Aims

This project builds on previous epigenetic studies in AD, with the aim to examine genome-wide alterations in 5mC and 5hmC independently in AD, and to profile the epigenetic landscape of the *ANK1* gene in AD in greater detail. The project has the following specific aims:

1. To perform the first EWAS to independently measure 5mC and 5hmC in AD brain using oxidative-bisulfite treatment in conjunction with the 450K Beadarray.
2. To investigate the specificity of 5mC changes in the *ANK1* gene to different neurodegenerative diseases using BS pyrosequencing.
3. To examine whether changes in histone modifications, namely H3K4me3 and H3K27me3, are seen in the *ANK1* gene in AD brain using Chromatin Immunoprecipitation quantitative PCR (ChIP-qPCR).

CHAPTER 2: MATERIALS AND METHODS

This chapter describes the generic materials and methods used throughout multiple chapters in this thesis. Detailed descriptions of materials and methods specific to particular chapters and statistical analysis performed is described in the relevant individual chapter. Much of the work presented in this chapter is based on an invited book chapter, which is currently under review (Smith 2018), which can be found in Appendix 2.

2.1. Human brain tissue cohorts

Brain tissue used in this thesis was collected from multiple Brain Banks within the UK. This included the LNDBB (http://www.kcl.ac.uk/iop/depts/cn/research/MRC-London-Neurodegenerative-Diseases-Brain-Bank/MRC-London-Neurodegenerative-Diseases-Brain-Bank.aspx), SWDBB, Oxford Brain Bank (http://www.medsci.ox.ac.uk/optima/information-for-patients-and-the-public/the-thomas-willis-oxford-brain-collection), Cambridge Brain Bank, Manchester Brain Bank and Newcastle Castle Brain Bank. From each cohort of samples we usually requested multiple brain regions, such as the EC, superior temporal gyrus (STG), CER, substantia nigra (SN) and striatum (STR). The use of these samples in the research presented in this thesis was approved by the University of Exeter Medical School Research Ethics Committee (reference number Apr14/C/041Δ1). Detailed information on each sample cohort is given in the relevant chapters.

2.2. Tissue dissection

Frozen post-mortem brain tissue was received on dry ice from the brain banks and then cut into 100mg sections for DNA extraction as follows. All stages of tissue dissection were performed in a sterile extraction hood and all materials were clean-room quality and autoclaved where possible, or purchased sterile/DNase-free. A sterile technique was used, and gloves were changed regularly.

1. The frozen brain tissue was placed in a sterile petri dish on dry ice and cut into small sections using a sterile scalpel
2. An empty frozen 1.5mL microcentrifuge tube was weighed and the weight was noted
3. A section of frozen tissue sample was transferred to the 1.5mL microcentrifuge tube with sterile tip or tweezers whilst on dry ice
4. The microcentrifuge tube was re-weighed and the brain weight was calculated (total weight-empty weight)
5. 100mg of tissue was used for DNA extraction

2.3. DNA extraction

This section documents in detail the nucleic acid extraction method used to isolate genomic DNA for studies described in Chapters 3 and 4. All plastic ware used during these procedures was sterile and DNase-free. Before starting experimental procedures, all the surfaces and materials were cleaned using ethanol and/or Alconox (Sigma-Aldrich, cat no.: 242985) to remove contaminants and DNases present.

2.3.1. Overview of genomic DNA isolation using phenol-chloroform

DNA extraction from brain tissue samples using a standard phenol-chloroform method has been widely utilised for EWAS using post-mortem brain tissue as it provides sufficient yields of high quality DNA for downstream purposes (De Jager et al. 2014, Lunnon et al. 2014). The method below is an adaptation of the phenol-chloroform method developed by Sambrook et al. (2006). Reagents used for the phenol-chloroform DNA extraction can be found in Tables 2.1, 2.2 and 2.3.

Name	Storage	pH	Final Concentration	Warnings	Supplier	Catalogue no.
NaCl	Room temp		75mM	Irritant	Sigma Aldrich	S3014
Tris-HCl	Room temp	pH8	10mM	Corrosive Toxic	VWR	733-1654
EDTA	Room temp	pH8	25mM	Irritant	Fisher Scientific	11568896
10% SDS	Room temp		$1/20^{th}$ final volume	Irritant Flammable	Sigma Aldrich	L3771
H_2O	Room temp			Not Hazardous		

Table 2.1; Reagents required to make lysis buffer.

Name	Storage	pH	Volume (mL)	Warnings	Supplier	Catalogue no.
1M tris-HCL	Room temp	pH8	10	Corrosive Toxic	VWR	733-1654
0.5M EDTA	Room temp	pH8	0.2	Irritant	Fisher Scientific	11568896
H_2O	Room temp		990	Not Hazardous		

Table 2.2; Reagents required to make 1x Te Solution (1L). *The solution was filtered with a 0.5 micron filter and then autoclaved before use.*

Name	Storage	Concentration	Warnings	Supplier	Catalogue no.
Proteinase K solution	-20°C	20mg/mL	Not Hazardous	Fisher Scientific	10172903
Ribonuclease A (from bovine pancreas)	-20°C	20µg/mL	Not Hazardous	Sigma Aldrich	R6513
Phenol / Chloroform / Isoamyl Alcohol (PCI)	4-8°C	100%	Toxic Flammable Irritant	Fisher Scientific	13148563
Chloroform ($CHCl_3$)	Room temp	100%	Irritant Toxic Flammable	Sigma Aldrich	C2432
Ethanol	-20°C	100%	Flammable	Sigma Aldrich	E7023
NaCl	Room temp	75mM	Irritant	Sigma Aldrich	S3014
Ethanol	Room temp	70%	Flammable	Sigma Aldrich	E7023
Elution Buffer (H_2O / Te)	Room temp	200-300µL per sample	Not Hazardous	See Table 2.2	

Table 2.3; Other reagents required for phenol-chloroform DNA extraction,

2.3.2. Lysis and digestion of cells

1. Fresh lysis buffer was prepared (1mL per 100mg tissue sample) (Table 2.3.1)
2. Four times weight/volume of lysis buffer was added to the tissue
3. The tissue was disrupted with a sterile disposable pestle and rotary pestle motor
4. Six times weight/volume of lysis buffer was added to the tissue
5. DNase-free RNase-A (Table 2.3) was added to a final concentration of 20µg/mL and samples were incubated for at least 1 hour at 37°C
6. Proteinase K (Table 2.3) was then added to each sample at a final concentration of 400µg/mL
7. The samples were incubated in a water bath at 50°C overnight
8. The following morning the level of sample lysis was checked. If the sample was not sufficiently lysed a p1000 pipette tip or a syringe was used to re-suspend the solution and break up any remaining tissue, being careful not to lose any of the solution
9. Proteinase K (Table 2.3) was added to each sample at a final concentration of 200µg/mL
10. The samples were left for at least one more hour in the water bath at 50°C
11. The water bath was then heated to 65°C and left for 20-30 minutes to deactivate the Proteinase K
12. The samples were then cooled to room temperature

2.3.3. Purification of DNA

1. Each sample was transferred to a 2mL phase lock tube and then an equal volume of PCI was added to the tube (Table 2.3)
2. The samples were mixed by inversion 20 times and centrifuged at 16,200 x g for 15 minutes
3. The aqueous layer was transferred into a new 2mL microcentrifuge tube
4. 1mL of 100% chloroform was added to each sample (Table 2.3)
5. The samples were mixed by inversion 20 times and centrifuged at 16,200 x g for 15 minutes
6. The aqueous layer was transferred into a new 2mL microcentrifuge tube

7. The chloroform extraction (steps 4 to 6) was repeated

2.3.4. Precipitation of DNA

1. 1 to 1.5mL of ice cold 100% ethanol (Table 2.3) was added to the sample and mixed by inversion
2. If DNA was not visible, 10-20µL 5M NaCl was added (Table 2.3)
3. The samples were left at -20°C overnight to further precipitate the DNA
4. The samples were centrifuged at 16,200 x g for 15 minutes
5. The supernatant was removed by pipetting, being careful not to dislodge the DNA pellet
6. A further 1mL of 70% ethanol (Table 2.3) was added to the DNA pellets
7. The samples were centrifuged for a further 10 minutes at 16,200 x g
8. The supernatant was removed by pipetting, being careful not to dislodge the DNA pellet. Pellets were left to air-dry for approximately 30 minutes
9. The samples were re-suspend in 200µL – 300µL H_2O
10. The pellet was dissolved overnight at 37°C

2.3.5. Determining the quality and quantity of isolated nucleic acids

1. A visual check was performed to make sure the DNA was well suspended
2. A 0.8% agarose gel was run for each sample to check for high molecular weight DNA, which is not degraded
3. 2µL of the DNA sample was also loaded on to a Spectrometer (NanoDrop) to determine the concentration and further check the quality of the DNA using 260/280 and 230/280 ratios (above 1.8 and ~2 respectively for pure DNA)

2.4. DNA bisulfite treatment (BS)

BS treatment, originally described by Frommer et al. (1992), is a method that provides a base-pair level sensitivity measure for DNA modifications (including 5mC). In this process un-modified cytosine is converted to uracil, but certain modifications to cytosine, including 5mC and 5hmC, are unreactive and remain as cytosine. The sequence can then be amplified via subsequent PCR, resulting in all uracil and thymine residues being amplified as thymine and only modified cytosine residues being amplified as cytosine. BS treatment using the Zymo EZ-96 DNA Methylation-Gold™ Kit (Cambridge Bioscience, cat no.: D5007) ideally utilises 500ng high quality DNA in 20µL to provide sufficient BS-treated material for profiling of cytosine modification using beadarrays or pyrosequencing. This method has been used to treat DNA samples in all EWAS studies of AD published to date (Bakulski et al. 2012, De Jager et al. 2014, Lunnon et al. 2014, Sanchez-Mut et al. 2014, Sanchez-Mut et al. 2016, Watson et al. 2016) and was used in Chapter 4 of this thesis. The method below is an adaptation of the sodium BS treatment method developed by the manufacturer. Steps 14-17 deviate from the original protocol which is available from http://www.zymoresearch.com/downloads/dl/file/id/59/d5007i.pdf.

2.4.1. Starting material

1. 500ng (25ng/µL in 20µL) of DNA was added to each well of a 96 well PCR plate

2.4.2. Sodium BS conversion

1. 130µL of the CT conversion reagent (Table 2.4) was added to each DNA sample in the conversion plate and mixed by pipetting
2. The plate was sealed with the provided film and transferred to a thermal cycler with the following steps performed:
    1. 98°C for 10 minutes
    2. 64°C for 2.5 hours
    3. 4°C storage for up to 20 hours

3. 400µL of M-Binding buffer was added to the wells of a Silicon-A™ binding plate mounted on a collection plate
4. The samples were transferred from the conversion plate (Step 2) to the corresponding wells of the Silicon-A™ binding plate (Step 3) and mixed by pipetting
5. The plate was centrifuged at 3,500 x g for 5 minutes then the flow-through was discarded
6. 400µL of M-Wash buffer (Table 2.5) was added to each well of the plate
7. The samples were centrifuged at 3,500 x g for 5 minutes
8. 200µL of M-Desulphonation buffer was added to each well and the plate was left to stand at room temperature (20-30°C) for 20 minutes
9. The plate was centrifuged at 3,500 x g for 5 minutes then the flow-through was discarded
10. 400µL of M-Wash buffer (Table 2.5) was added to each well of the plate
11. The plate was centrifuged at 3,500 x g for 5 minutes then the flow-through was discarded
12. 400µL of M-Wash buffer (Table 2.5) was added and centrifuged for 10 minutes
13. The Silicon-A™ binding plate was placed onto an elution plate
14. 15µL of M-Elution buffer was added directly to each well of the Silicon-A™ binding plate.
15. The plate was incubated for 5 minutes at room temperature
16. The plate was then centrifuged at 4,000 x g for 3 minutes to elute the DNA
17. Steps 15 to 17 were repeated

The eluted DNA was ready for immediate analysis or was stored at or below -20°C for later use. For long term storage, BS treated DNA was stored at or below -70°C.

2.4.3. Quality control (QC)

BS conversion quality was assessed using BS-specific PCR amplification (Section 2.7.1) followed by gel electrophoresis (Section 2.7.2). A single visible PCR band was taken as an indication of suitable BS conversion efficiency.

Name	Volume (mL)	Supplier
M-Dissolving Buffer	0.5	Zymo
M-Dilution Buffer	3	Zymo
CT Conversion Reagent	Bottle provided	Zymo
Ultra-pure H_2O	9	User

Table 2.4; CT Conversion Reagent required for BS conversion. using the Zymo EZ-96 DNA Methylation-Gold™ Kit (cat no.: D5007). *The CT conversion reagent is supplied as a powder and must be dissolved prior to use and mixed at room temperature with frequent vortexing or shaking for 15 minutes*

Name	Volume (mL)	Supplier
M-Wash Buffer	36	Zymo
100% Ethanol	144	User

Table 2.5; Reagents required for the M-Wash Buffer for BS conversion using the Zymo EZ-96 DNA Methylation-Gold™ Kit (cat no.: D5007).

2.5. DNA methylation and hydroxymethylation profiling in parallel

Oxidative BS (OxBS) treatment is one of the first methods enabling the quantitative mapping of 5mC in the absence of confounding by 5hmC in genomic DNA at single-nucleotide resolution. Using a selective chemical oxidation of 5hmC to 5fC, followed by BS conversion, results in the conversion of 5hmC to uracil. As such only cytosines that are methylated are protected from the conversion and after subsequent PCR amplification, one can assess the level of just 5mC. Simultaneous BS treatment of a matched sample of DNA, provides a value for the total DNA modifications (5mC + 5hmC) at each site, as both modifications are protected from the conversion and are not converted to uracil (Booth et al. 2012). Subsequently after PCR amplification, the subtraction of the OxBS signal from the standard BS signal gives a 5hmC proxy value for each nucleotide (Figure 2.1.).

OxBS treatment using the CEGX® TrueMethyl® Array kit requires 1µg high molecular weight genomic DNA per sample. This is sufficient to allow 500ng to be processed through each of the BS and OxBS conversion processes in parallel. The protocol described below results in a final volume of 12µL of BS treated DNA and 12µL of OxBS treated DNA for downstream analyses. The method below was used in Chapter 3 of this thesis as per the manufacturer's protocol, which is available from https://www.cambridge-epigenetix.com/resources/user-guides.

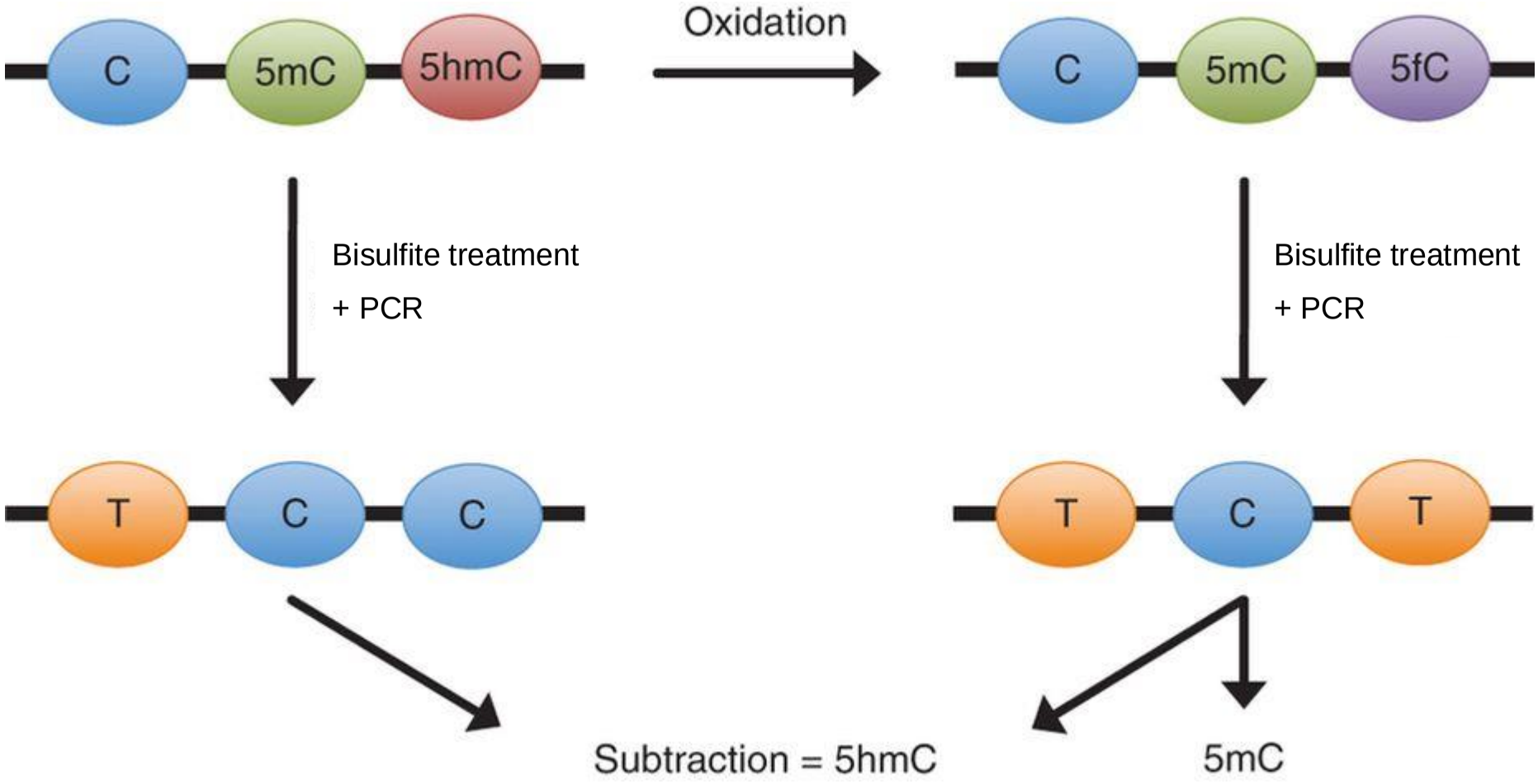


Figure 2.1; Oxidative bisulfite conversion. (Booth et al. 2013)

2.5.1. Starting Material

1. The digestion control sample was thawed on ice
2. 1µg DNA sample was added to a 1.5 mL microcentrifuge tube and the final volume was adjusted to 49 µL using ultra-pure H_2O from the TrueMethyl® kit
3. 1µL of cutting control was added to each DNA sample and mixed well before vortexing

2.5.2. CEGX TrueMethyl® conversion

2.5.2.1. DNA oxidation

1. Before starting the workflow the heat block was set to 37°C and magnetic binding solution 1 (Table 2.6) and 80% acetonitrile (Table 2.7) was prepared
2. The magnetic bead binding solution 1 was vortexed thoroughly before use to ensure the solution is homogenous and 100µL was added to each 1.5 mL microcentrifuge tube containing 50µL DNA and vortexed to mix
3. The samples were incubated for 20 minutes at room temperature
4. The samples were briefly centrifuged to collect the sample at the bottom of the tube
5. The tubes were placed into a magnetic separation rack to pellet beads for 5 minutes at room temperature
6. The supernatant was carefully removed and discarded. If beads were accidentally pipetted they were dispensed back into the tube and steps 5 and 6 were repeated
7. The following wash steps were performed, leaving the 1.5mL tubes in the magnetic separation rack:
    a. 1mL 80% acetonitrile (Table 2.7) wash was added to the tubes without disturbing the bead pellet
    b. The 1mL 80% acetonitrile wash was removed and discarded, carefully avoiding aspiration of the bead pellet
    c. Steps 8a and 8b were repeated another two times so that ultimately three 1ml 80% acetonitrile washes were performed in total. As much of the final wash was removed by pipetting as possible

d. The bead pellets were air dried for 5 minutes at room temperature, with the lids of the tubes left open

8. The following denaturation steps were then completed:
    a. The 1.5mL tubes were left in the magnetic separation rack, and 50µL of denaturing solution was added directly onto the bead pellet. The lids of the tubes were closed before removing them from the magnetic separation rack
    b. The denaturing solution was vortexed thoroughly to fully suspend the bead pellet
    c. The sample was incubated for 5 minutes at room temperature
    d. The sample was incubated for 30 minutes at 37°C (performed in the pre-heated benchtop heating block)
9. The tubes were briefly centrifuged to collect the beads at the bottom of the tubes
10. The tubes were placed into a magnetic separation rack and beads were pelleted for 5 minutes at room temperature
11. Avoiding the bead pellet, 50µl eluate was transferred into a new 1.5mL tube and centrifuged briefly. The sample was placed on ice before proceeding immediately to step 12
12. The DNA oxidation reaction was performed in 0.2mL PCR tubes as described in Table 2.8.
13. The oxidation reaction mix was incubated at 40°C for 30 minutes in a PCR thermocycler with a heated lid
14. The samples were centrifuged at 14,000 x g for 10 minutes at room temperature to pellet any black precipitate
15. The orange supernatant was transferred into new 0.2mL PCR tubes and this was used in subsequent steps of the protocol

Name	Volume (mL)	Supplier
Magnetic Bead Solution	0.24	CEGX
Binding Buffer 1	12	CEGX

Table 2.6; Magnetic Bead Binding Solution 1 required for the CEGX TrueMethyl® Array kit. *The Magnetic Bead Binding Solution 1 (green cap) is prepared by the addition of Binding Buffer 1 to pre-vortexed Magnetic Bead solution. After adding the buffer, the green cap was replaced and vortex thoroughly. The Magnetic Bead Binding Solution 1 was stored between 2-8°C. The solution was equilibrated to room temperature for a minimum of 30 minutes before use.*

Name	Volume (mL)	Supplier
100% Acetonitrile	32	User
ultra-pure H_2O	8	CEGX

Table 2.7; 80% Acetonitrile required for the CEGX TrueMethyl® Array kit*. This was prepared fresh for each experiment and mixed by inversion. The amounts listed below are sufficient for 12 samples.*

	Volume per reaction (µL)	
Component	**oxBS**	**BS**
Denatured DNA from step 12	24	24
Oxidant Solution	1	-
Ultra-Pure Water	-	1
Total volume	25	25

Table 2.8; Reagents required for the Oxidation reaction in the CEGX TrueMethyl® Array kit. *Oxidant Solution was thawed on ice prior to use. The samples were mixed briefly by vortexing, and then briefly centrifuged before placing in a PCR thermocycler.*

2.5.2.2. BS conversion

1. BS reagent solution was prepared by adding 4.2mL BS diluent to one BS reagent aliquot (sufficient for 24 reactions). The lid was tightly sealed and the solution was incubated at 60°C for 15 minutes (using a water bath). Subsequently the samples were vortexed until the BS reagent solution was completely dissolved. It was ensured that the same BS reagent solution aliquot was used for each OxBS and BS treated sample pair
2. The oxidised samples from step 15 of the DNA oxidation protocol were equilibrated to room temperature
3. The BS conversion reaction setup was performed in 0.2mL PCR tubes as described in Table 2.9.
4. The 0.2mL PCR tubes containing the BS conversion reaction mix were placed into a thermocycler. The BS DNA conversion was performed using the thermal profile shown in Table 2.10.

Component	Volume per reaction (µL)
DNA solution (step 15)	25
BS Reagent Solution	170
BS Additive	5

Table 2.9; Reagents required for the BS conversion mix in the CEGX TrueMethyl® Array kit. *To this sample tube the BS reagent solution was first added, followed by the BS additive. The PCR tubes were then vortexed thoroughly and briefly centrifuged to collect the samples at the bottom of the tubes.*

Step	Time (min)	Temperature (°C)
Denaturation	5	95
Incubation	20	60
Denaturation	5	95
Incubation	40	60
Denaturation	5	95
Incubation	165	60
Denaturation	5	95
Incubation	20	60
Denaturation	5	95
Incubation	40	60
Denaturation	5	95
Incubation	165	60
Hold	Indefinite	20

Table 2.10; The BS conversion thermal profile for the CEGX TrueMethyl® Array kit.

2.5.2.3. Clean up of BS converted DNA

1. The magnetic binding solution 2 was prepared (Table 2.11.)
2. The magnetic bead binding solution 2 and desulfonation buffer (Table 2.12.) were equilibrated to room temperature for a minimum of 30 minutes before use. A fresh stock of 70% ethanol was also prepared for the experiment (Table 2.13.)
3. Once the BS conversion was complete, tubes were centrifuged at 14,000 x g for 10 minutes at room temperature. This pelleted any precipitated salts accumulated during the BS incubation
4. Without disturbing the pellet, 195µL of the supernatant was transferred to a new 1.5mL microcentrifuge tube
5. The magnetic bead binding solution 2 (Table 2.11) was vortexed thoroughly to ensure the solution was homogenous before aliquoting
6. 1mL of magnetic bead binding solution 2 (Table 2.11) was added to each of the 1.5mL microcentrifuge tubes containing 195µL BS converted sample from step 23. The lid was closed, tube inverted and the base flicked to dislodge the sample, before being vortexed to mix
7. The samples were incubated for 20 minutes at room temperature
8. At the end of the incubation, samples were briefly centrifuged to collect the sample at the bottom of the tube
9. The samples were then placed into a magnetic separation rack to pellet beads for 30 minutes at room temperature
10. The supernatant was removed and discarded, avoiding the bead pellet. If beads were accidentally pipetted, they were dispensed back into the tube and steps 9 and 10 were repeated
11. The following wash steps were completed, leaving the 1.5mL tubes in the magnetic separation rack:
    a. 1.4mL 70% ethanol wash (Table 2.13) was added to the tubes without disturbing the bead pellet
    b. The 1.4mL 70% ethanol wash was removed and discarded, carefully avoiding aspiration of the bead pellet. As much of the wash was removed as possible
12. The following desulfonation steps were then completed:

a. The 1.5mL tubes were left in the magnetic separation rack and 1mL Desulfonation buffer (Table 2.12) was added directly onto the bead pellet. The lids of the tubes were closed before removing from the magnetic separation rack. The tubes were vortexed to fully suspend the bead pellet in desulfonation buffer and then centrifuged briefly
b. The tubes were placed back into the magnetic separation rack and incubated for 5 minutes at room temperature in the desulfonation buffer to pellet the beads
c. The 1ml desulfonation buffer was removed and discarded, carefully avoiding aspiration of the bead pellet. As much of the desulfonation buffer was removed by pipetting as possible

13. The following wash steps were completed, leaving the 1.5mL tubes in the magnetic separation rack:
    a. 1mL 70% ethanol wash was added to the tubes without disturbing the bead pellet
    b. The 1mL 70% ethanol wash was removed and discarded, carefully avoiding aspiration of the bead pellet
    c. Steps 13a and 13b were repeated so that ultimately 1mL 70% ethanol washes were performed two times in total. As much of the final wash was removed as possible
    d. The bead pellets were air dried for 15 minutes at room temperature, by leaving the lids of the tubes open
14. Following the wash steps the following elution steps were performed:
    a. 12µL of elution buffer (provided) was added directly onto the bead pellet, leaving the 1.5mL tubes in the magnetic separation rack. The lids of the tubes were closed before removing from the magnetic separation rack
    b. The bead pellet was resuspended in the elution buffer by thorough vortexting. The sample was incubated for 20 minutes at room temperature to elute the TrueMethyl® converted DNA from the beads
15. At the end of the incubation (step 14b), the 1.5mL microcentrifuge tubes were briefly centrifuged to collect the sample at the bottom of the tube. The tubes were then placed into a magnetic separation rack to pellet beads for 5 minutes at room temperature

16. 12µL eluate was transferred into a new 1.5mL tube, avoiding the bead pellet

The DNA was then ready for immediate analysis or could be stored at or below -20°C for later use. For long term storage, the DNA was stored at -80°C.

Name	Volume (mL)	Supplier
Magnetic Bead Solution	2.64	CEGX
Binding Buffer 2	220	CEGX

Table 2.11; Preparation of Magnetic Bead Binding Solution 2 for the CEGX® TrueMethyl® Array kit. *The Magnetic Bead Binding Solution 2 was prepared by the addition of pre-vortexed Magnetic Bead solution to Binding Buffer 2. Subsequently the solution was vortexed thoroughly. Magnetic Bead Binding Solution 2 was stored between 2-8°C.*

Name	Volume (mL)	Supplier
Desulfonation Buffer	Bottle provided	CEGX
100% Ethanol	150	User

Table 2.12; Preparation of Desulfonation Buffer for the CEGX TrueMethyl® Array kit. *The Desulfonation Buffer was diluted with 100% Ethanol before use with the lid sealed tightly and mixed thoroughly by inversion. Desulfonation Buffer is stored long term between 2-8°C to prevent ethanol evaporation.*

Name	Volume (mL)	Supplier
100% Ethanol	70	User
ultra-pure H2O	30	CEGX

Table 2.13; Preparation of 70% Ethanol for the CEGX TrueMethyl® Array kit. *The amounts below are sufficient for 12 samples. The ethanol was thoroughly mixed by inversion.*

2.5.3. Quality control

To assess OxBS conversion quality it was recommended to integrate a Digestion control (provided) with the initial 1µg gDNA sample in Section 2.5.1 step 3 and interrogate the control using PCR, followed by gel electrophoresis, as described below.

2.5.3.1. Amplification of digestion control

1. A PCR reaction mix was prepared as shown in Table 2.14.
2. Samples were then amplified according to the thermocycling conditions outlined in Table 2.15.

2.5.3.2. Taq$^{\alpha}$I digestion of the digestion control amplicon

1. The following digestion reaction mix was prepared as outlined in Table 2.16.
2. Taq$^{\alpha}$I digestion mixes were incubated for 18 hours at 65°C
3. Samples were then denatured at 80°C for 20 minutes
4. Digestion mixes were then run on a 2% agarose gel (Section 2.7.)

Component	Volume (µL)
Ultra-Pure Water	39
10X polymerase buffer	5
10mM dNTP mix	2
Digestion Control Fwd PCR Primer (100µM)	0.5
Digestion Control Rev PCR Primer (100µM)	0.5
DNA (from 2.5.2 step 35)	1
Uracil tolerant DNA polymerase (5U)	1
TOTAL volume	50

Table 2.14; Digestion control PCR mix for assessing the conversion quality of oxidative bisulfite (OxBS) treated DNA.

Segment	Number of Cycles	Temperature (°C)	Duration (seconds)
1	1	95	120
2	40	95	30
		60	30
		72	15
3	1	4	∞

Table 2.15; PCR conditions for the digestion control.

Component	Volume (µL)		
	Digestion Control	Digestion –ve Control	Cutting Control
10X Taq$^{\alpha}$I	2.2	2.2	2.2
Taq$^{\alpha}$I RE (20U/µL)	2	-	2
100X BSA	4	4	4
Processed Digestion Control amplicon DNA (10ng/µL)	10	10	-
Cutting Control DNA (20ng/µL)	-	-	5
Ultra-Pure Water	3.8	5.8	8.8
TOTAL volume	22	22	22

Table 2.16; Taq$^{\alpha}$I digestion conditions.

2.6. Polymerase chain reaction (PCR)

The amplification of DNA is used in every chapter within this thesis. PCR is a method used extensively to amplify a segment of DNA to generate numerous copies of the same fragment. Various components, including DNA, nucleotides, buffers, primers and polymerase are combined together and subjected to cycles of heating and cooling in a thermocycler. During the first stage of PCR the mix is heated to a high temperature to activate the heat sensitive polymerase taq. This is in turn followed by three steps:

1. A denaturation step, where the mix is heated to 95ºC to denature the double-stranded DNA
2. An annealing step where the mix is cooled to a primer-specific temperature (usually between 50ºC and 65ºC) to allow the primers to anneal with high specificity to the correct annealing sequence in the DNA
3. An extension step at 72ºC to allow the taq polymerase to synthesise the complementary strand of DNA using the deoxynucleotides (dNTPs)

These three steps are repeated for a specific number of cycles to allow the synthesis of an exponential number of DNA amplicons. A final step of 72ºC is added to allow a final extension.

The PCR in Chapters 3 and 4 in this thesis was performed using BS-converted genomic DNA as the reaction template (see Section 2.4 and 2.5). This presents more of a challenge than using standard unconverted genomic DNA for several reasons. First, extended incubations using BS-treated DNA can lead to extensive damage to the DNA template. Second, BS treatment leads to reduced sequence complexity (the DNA largely comprises three bases rather than four), leading to a higher redundancy of the target sequence, and third, regions of interest often lie within CG-rich sequences, which become long stretches of Thymine following BS conversion, which can cause polymerase slippage. These reasons result in a higher chance of mispriming and non-specific PCR amplification. These issues were addressed in this thesis with careful primer design, using the following criteria which are widely reported to be optimal for successful BS-PCR:

1. Primers did not contain any CpG sites within their sequence to avoid discrimination between methylated or unmethylated DNA
2. Primers were not placed in a repetitive region
3. Primers only contained non-CpG cytosines within their sequence to ensure exclusive amplification of BS-converted DNA
4. Primer length was at least 20 bases, to reduce non-specific amplification

The cycling conditions, primer sequences and reagents used in the *ANK1* PCR reactions in this thesis (Chapters 3 and 4), together with their description and quantities in a standard PCR reaction, are described in Table 2.17, Table 2.18 and Table 2.19, respectively.

Step	Temperature (°C)	Time	Number of Cycles
Hotstart	95	15min	1
Denaturation	95	45sec	40
Annealing	60	45sec	
Extension	72	3min	
Final Extension	72	10min	1

Table 2.17; ANK1 Polymerase chain reaction (PCR) thermocycling conditions.

Primer Name	Sequence	Length (bp)	GC %
ANK1_F	GTGGAGGGAGGTGTTATGTA	20	50
ANK1_R	[Btn]CTTCCCTTCCTTTCCTAAATCATT	24	37.5
ANK1_CPG1_S	TTGGGTTTTTTAGGGT	16	37.5
ANK1_CPG2_S	AGGGGTTTTTTTATAGTTATT	21	23.8

Table 2.18; ANK1 Polymerase chain reaction (PCR) and pyrosequencing primer sequences.

Reagent	Function	Reagent Concentration	Volume (µL)	Supplier	Catalogue no.
PCR Buffer	maintains optimal pH	10X	3	Solis BioDyne	01-02-01000
Magnesium Chloride ($MgCl_2$)	Required substrate for the Taq polymerase	25mM	1.98	Solis BioDyne	01-02-01000
DNA nucleotides	DNA bases needed for the synthesis of new DNA	10mM	0.3	Fisher Scientific	1183-3933
PCR primers	Short, single-stranded oligonucleotides complementary to target sequence	10µM	1.5	IDT	
Taq DNA Polymerase	Heat resistant enzyme that catalyses the PCR reaction	5U/µL	0.3	Solis BioDyne	01-02-01000
DNA	Single stranded BS or OxBS treated DNA	10ng/µL	2		
Ultra-pure Water	Makes up reaction volume	Final volume of 30µL	20.92		

Table 2.19; ANK1 Polymerase chain reaction (PCR) reagents and volumes.

2.7. Agarose gel electrophoresis

The visualisation of DNA/Chromatin via gel electrophoresis was needed in various quality control stages in each data chapter, for example checking the efficiency of BS and/or OxBS conversion in Chapters 3 and 4, and checking the level of chromatin shearing in Chapter 5. In addition gel electrophoresis was used to check the quality of the all DNA extracted as described in Section 2.3.

Agarose gel electrophoresis enables the separation of DNA molecules based on their size. It is most commonly used to assess the quality and quantity of DNA and to determine the efficiency of molecular biology techniques such as DNA extraction or PCR amplification. The gels used in this thesis ranged from 0.8 to 2.0% agarose (Sigma Aldrich, cat no.: A9539) in 1% tris-borate EDTA (TBE) buffer (Fisher Scientific, cat no.: 10031223). The concentration of agarose influences the amount of resistance applied to molecules passing through the matrix, therefore the concentration of agarose selected is dependent on the size of the molecule to be separated. A small amount of the Syto60 compound (Fisher Scientific, cat no.: 10194852) is added ubiquitously to the gel before it cools. Once cooled DNA samples were added to the gel matrix, an electrical charge was applied across it, causing the negatively charged DNA to migrate towards the positive anode, at a speed determined by the DNA fragment size. Smaller molecules move through the matrix faster, and therefore move further through the gel in the allocated time. The separated DNA was then viewed using the fluorescent Syto60, which intercalates into the DNA structure allowing it to be visualised after laser exposure. The laser light is absorbed by the Syto60 and reemitted; these wavelengths of light are then detected by the Odyssy Clx system (LI-COR, USA), which were then observed. In this thesis, agarose gel electrophoresis was used for the inspection of DNA extraction quality (Chapter 3 and 4) using a 0.8% gel, PCR products (Chapter 3 and 4) using a 2% gel and visualising sheared Chromatin (Chapter 5) using a 1.5% gel.

2.8. Infinium HumanMethylation450 BeadChip Array

The 450K array combines BS/OxBS converted DNA and whole-genome amplification with direct, array-based capture and scoring of the CpG loci (Pidsley et al. 2016). This enables the quantification of 5mC levels at 485,577 CpG sites, covering 99% of Reference Sequence database (RefSeq) genes. Important regulatory regions such as CpG islands (96% covered), island shores and shelves, 5′ and 3′ UTRs, promoters and gene bodies are also covered (Bibikova et al. 2011). Signal intensity is measured with the Illumina iScan or HiScan system to generate beta values, a measure of the levels of DNA modification, at each locus. One or two probes are used to interrogate a single CpG locus, depending on the probe design for a particular CpG site, as the Infinium I design has two probes per site, whilst the Infinium II has one probe per site. The type I probe design is based on the assumption that the methylation status of CpG sites within 50bp of the probe are correlated with the query CpG (Illumina 2015). However, if underlying probes are not associated with 5mC at the target site this could lead to a source of bias. The Infinium II probes use a single bead type with the 5mC state determined by a single base extension step after hybridisation by two different coloured dyes (green for methylated/modified sites and red for unmethylated/unmodified sites). This design includes the addition of a degenerate R [A/G] base at underlying CpG sites, and therefore is less influenced by nearby co-methylation patterns. Allele-specific single base extension of the probes incorporates a biotin nucleotide or a dinitrophenyl labelled nucleotide. Signal amplification of the incorporated label further improves the overall signal-to-noise ratio of the assay. This method was used in Chapter 3.

A full description of this methodology, used in Chapter 3, can be found on the Illumina website http://www.illumina.com/products/by-type/microarray-kits/infinium-methylation-epic.html. Briefly, the first stage in the process includes a denaturing and neutralisation stage before the BS or OxBS treated DNA is whole-genome amplified. The denatured DNA was then isothermally amplified in an overnight step. The whole-genome amplification uniformly increases the amount of the DNA sample by several thousand folds without significant amplification bias. A controlled enzymatic process then fragmented the amplified product. The process uses endpoint fragmentation to prevent over fragmentation. After an isopropanol

precipitation, centrifugation at 4°C collected the fragmented DNA. The precipitated DNA was then suspended in Hybridisation buffer. DNA samples were then dispensed onto BeadChip arrays and incubated in an Illumina Hybridisation Oven to hybridise the samples onto the BeadChip array surface. 12 samples were applied to each BeadChip array, which were kept separate with an IntelliHyb seal. DNA samples annealed to locus-specific sites on the BeadChip array. Un-hybridised and nonspecifically hybridised DNA was washed away and the BeadChip array then underwent extension and staining in capillary flow-through chambers. Single-base extension of the oligonucleotides on the BeadChip array, using the captured DNA as a template, incorporated detectable labels on the BeadChip array and determined the 5mC/5hmC level at the specific CpG sites. Finally, the BeadChip array was imaged on the Illumina HiScan or iScan System (Illumina, San Diego, CA, USA), using a laser to excite the fluorophore of the single-base extension product on the beads. The scanner recorded high resolution images of the light emitted from the fluorophores. These in turn were used by the genome studio® software to produce a beta value for each locus.

2.8.1. Data normalisation and quality control

The DNA modification level at each CpG site was determined by calculating the ratio of the intensity of the fluorescent signal for M (modified) and U (unmodified), which gives a β (DNA modification) value for each site ranging from 0 (*i.e.* all cytosines at that site are unmodified) to 1 (*i.e.* all cytosines at that site are modified). The β value is calculated by the following equation:

$$\beta = \frac{Intensity\ M}{Intensity\ M + Intensity\ U + 100}$$

To avoid negative values after background adjustment, Illumina recommends adding a constant offset (+100) to the denominator to regularise β values when both methylated and unmethylated probe intensities are low (Du et al. 2010). The Illumina 450K array has become one of the most popular methodologies to assess genome-wide DNA modifications in AD in recent years (De Jager et al. 2014, Lunnon et al. 2014, Watson et al. 2016), and as such was selected for the EWAS presented in

Chapter 3 of this thesis. However, although this technology has been proven to be robust, it is not without its own limitations: for example, one disadvantage is the difference in performance between the two types of probe (Dedeurwaerder et al. 2011). Specifically, the β-values obtained from Infinium II probes have been shown to be less accurate and reproducible than those obtained from Infinium I probes. As a result this has led to the development of several normalisation methods (Wang et al. 2015). Our research group has developed the *wateRmelon* package in R 3.3.2 (Vienna 2012, Pidsley et al. 2013), which offers a range of normalisation function and tools that were used to pre-process and normalise the DNA modification data presented in this thesis.

The presence of SNP variation within close proximity of the query CpG site is another issue with the probe design on the array. The 5mC data at these sites could be confounded by the presence of the polymorphism (Chen et al. 2013, Price et al. 2013). Therefore the DNA modification levels detected by these probes could simply be a representation of the underlying genetic polymorphisms. Additionally, 6-8% of the probes in the array have been found to cross-hybridise with other genomic locations and therefore do not accurately estimate the 5mC levels at the annotated site (Chen et al. 2013, Price et al. 2013). As a result of these issues, probes known to either cross-hybridise with another site or contain SNP variation were excluded prior to further analysis in this thesis.

2.9. Pyrosequencing

Pyrosequencing was used in both Chapters 3 and 4 of this thesis. Pyrosequencing is designed to detect changes in specified variable positions in, or base-calling of, DNA prepared from biological samples in molecular biology applications. The principle of pyrosequencing is based on work by (Nyren 1987) who noted that DNA polymerisation can be monitored by measuring pyrophosphate production, which in turn can be detected as light. Pyrosequencing, although originally designed as a less complex method of sequencing compared to Sanger sequencing, is now also widely used to sequence BS treated DNA and therefore gain highly sensitive measures of DNA modification at CpG sites. The fundamental basis of pyrosequencing is that pyrophosphate is released when a deoxyribonucleotide triphosphate is added to the end of a growing strand of DNA. The pyrophosphate is then converted by a sulfurylase enzyme to adenosine triphosphate, a cofactor for the enzyme luciferase. Luciferase oxidises the oxidation of luciferin to oxyluciferin and light. As a result of deoxyribonucleotide triphosphates being sequentially added to the reaction and because the light emitted is continuously monitored, the DNA sequence and therefore the level of DNA modification at a specific genomic locations can be determined (Harrington et al. 2013) (Figure 2.1). Pyrosequencing was used as a validation method in Chapter 3 and the main method of interrogation in Chapter 4. Biotinylated PCR primers (Table 2.18) were used to amplify a region of DNA for pyrosequencing. In addition, sequencing primers specific to the region of interest (approximately 100bp) were designed, these primers also had stringent design conditions as follows:

1. Sequencing primers did not contain any CpG sites within their sequence to avoid discrimination between methylated or unmethylated DNA
2. Sequencing primers were not placed in a repetitive region
3. Sequencing primers only contained non-CpG cytosines within their sequence to ensure exclusive amplification of BS-converted DNA
4. Sequencing primer length was no longer than 21 bases, to ensure efficient annealing
5. Sequencing primer annealing temperature was designed to be as close to 40°C as possible to be efficient within the pyrosequencer

All primers for PCR and pyrosequencing where designed using the PyroMark Assay Design software from Qiagen. CpG assay designs were created using the PyroMark Q24 software from Qiagen, using standard conditions. Two BS control regions were used per assayed region to confirm BS conversion efficiency. The method below details the process of pyrosequencing as per manufacturer's instructions on a targeted genomic region of BS or OxBS treated DNA as used in this thesis. A full guide is also available from the manufacturer (https://www.qiagen.com/gb/resources/resourcedetail?id=59f0275d-e60f-4517-b786-b0e0ca13952e&lang=en). For further practical information please see https://www.youtube.com/watch?v=LbEz9GBKW-s .

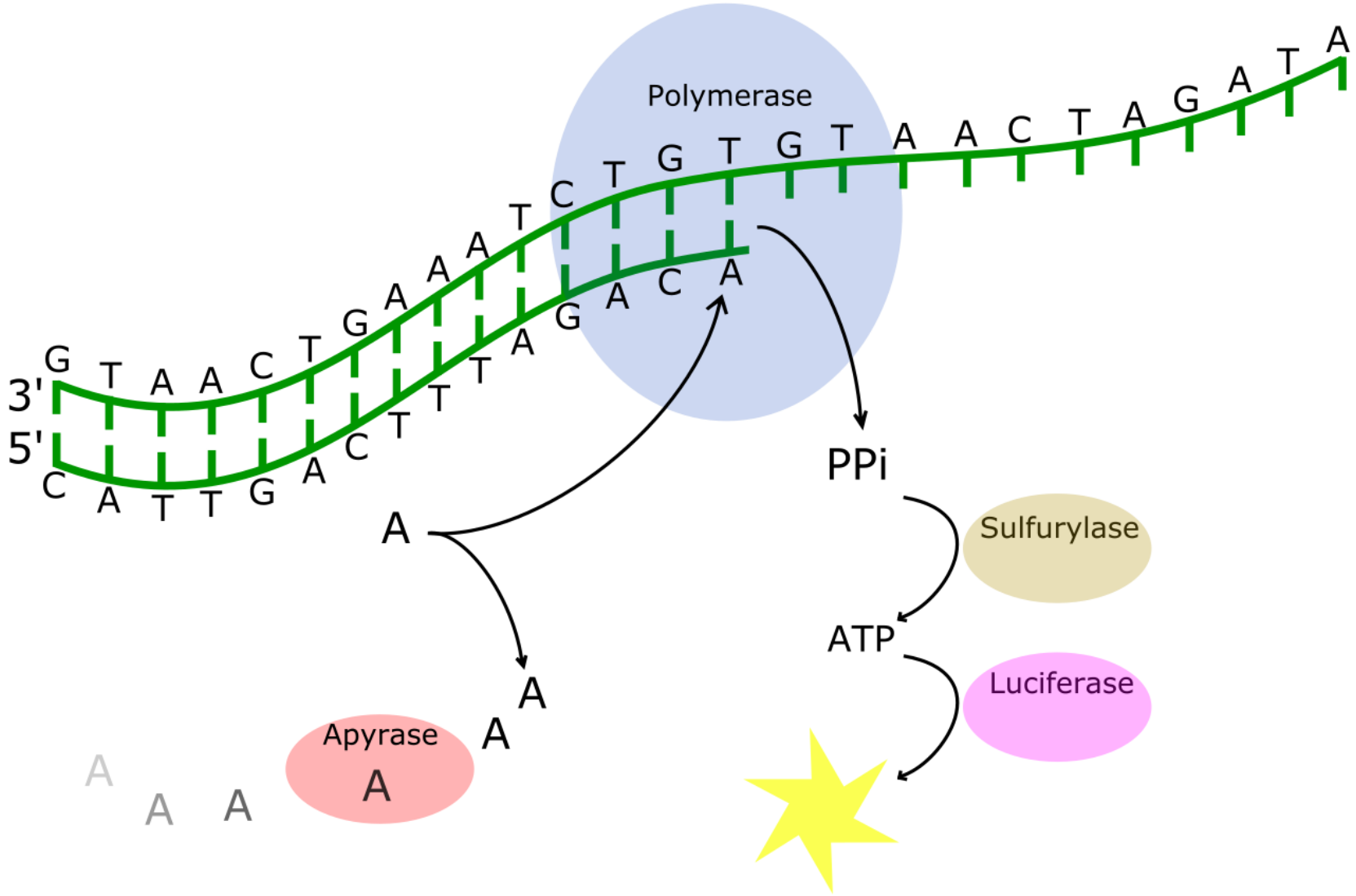


Figure 2.2; Chemical reaction of pyrosequencing. *Emitted light is proportional to the number of bases incorporated into the sequence.*

2.9.1. Immobilising PCR product

1. Sepharose beads (Streptavidin Sepharose High Performance, GE Healthcare, cat no.:11565015) were gently shaken to re-suspend
2. 2µL of beads (per sample) and Binding Buffer (Qiagen, cat no.: 979306) (40µL per sample) were combined in a 1.5mL centrifuge tube
3. Ultra-pure H_2O was added to make a total volume of 60µL per sample
4. 60µL of the immobilising solution was pipetted into each well of a 24-well PCR plate
5. 20µL of biotinylated PCR product was added to each well of the PCR plate
6. The plate was then sealed and agitated for at least 10 minutes (1,400rpm)

2.9.2. Preparing the vacuum workstation

1. The appropriate troughs were filled with each reagent as follows:
    a. 50mL 70% Ethanol (Sigma Aldrich, cat no.: E7023)
    b. 50mL Denaturation Solution (0.2M Sodium Hydroxide (NaOH)) (Qiagen, cat no.: 979307)
    c. 70mL 1 X Wash Buffer (Qiagen, cat no.: 979308)
    d. 70mL high-purity water
    e. 70mL high-purity water
2. The vacuum pump was switched on and filter probes were lowered into trough (e) (high-purity water). All water was washed through the probes and the trough was re-filled.
3. The vacuum pump was turned off and returned to the “parking” position

2.9.3. Sequencing primer setup

1. The sequencing primer was diluted to 0.3µM in Annealing Buffer (Qiagen, cat no.: 979309); 25µL of the resulting solution was pipetted into each well of a Q24 well plate

2.9.4. Combining the Q24 plate

1. Once the PCR plate has finished its agitation it was moved to the Vacuum Workstation
2. The vacuum tool was turned on and the filter probes were immersed in the wells of the PCR pate
3. The vacuum tool with the DNA bound sepharose beads was then moved to the ethanol trough (a) and ethanol was aspirated for 5 seconds
4. The vacuum tool was then moved to the denaturation solution trough (b) and denaturation solution was aspirated for 5 seconds
5. The vacuum tool was then moved to the wash buffer trough (c) and wash buffer was aspirated for 10 seconds
6. The vacuum tool was then raised to 90° and the filter probes were left to dry for 20 seconds
7. The vacuum tool was then turned off and lowered into the waiting Q24 plate containing the sequencing primer
8. The vacuum tool was then agitated from side to side to dislodge the beads into the solution
9. The Q24 plate was then transferred to a heat block and incubated at 80°C for 2 minutes. It was then left to stand at room temperature for 5 minutes

2.9.5. Loading the pyrosequencer

1. The sequencing cartridge was filled with reagents as described for each assay in the pre-run information produced by the Pyromark Q24 software
2. The cartridge and plate were loaded into the sequencer; the appropriate run was loaded from an external memory drive and was run

2.9.6. Pyrosequencing data quality control

Analysis of the pyrogram was performed using the PyroMark Q24 software. BS conversion efficiency of less than 90% resulted in the exclusion of the affected samples. Using the inbuilt quality checker, traces were included or excluded from downstream analysis. 5mC/5hmC values were reported as percentages of modifications at each CpG site and these were taken forward for downstream analysis.

2.10. Chromatin Immuno-Precipitation (ChIP)

This protocol was utilised in Chapter 5. The ChIP PCR method used in this thesis enables the relative quantification of specific histone modifications (H3K4me3 and H3K27me3) at specific loci within the *ANK1* gene.

2.10.1. ChIP protocol

ChIP uses immunoprecipitation to investigate the interaction between histone proteins and the DNA in a cell. It is used to determine whether specific histone proteins/ histone modifications are associated with specific genomic regions or loci (Orlando 2000). In the ChIP method DNA and associated proteins on chromatin in living cells or tissues are crosslinked using formaldehyde. The DNA-protein complexes (chromatin) are then sheared to between 100-600bp in length by sonication. Cross-linked DNA fragments associated with the histone modification(s) of interest are then selectively immunoprecipitated from the cell debris using a protein modification-specific antibody. The associated DNA fragments are then purified and quantified via quantitative PCR (qPCR). The reagents used were provided in the True MicroChIP Kit (Diagenode, cat no.: C01010130).

2.10.1.1. Tissue preparation

1. Tissue was cut into sections of between 10-15mg as described in Section 2.2.

2.10.1.2. DNA-protein cross linking

1. 5472.5µL of 1X phosphate buffered saline (PBS) (Sigma Aldrich cat no.: P5493) and 27.5µL of protease inhibitor cocktail (PIC) (Diagenode, cat no.: C01010130) were combined. 1mL of cold PBS plus PIC was added to each sample and tissue was disaggregated using a dounce homogeniser (30-50 times) to get a homogenous suspension. Samples were transferred into 1.5 mL LoBind Eppendorf tubes. The dounce homogeniser was washed with water, ethanol and then water again between samples

2. 27µL of 36.5% formaldehyde (Sigma Aldrich cat no.: F8775) was added to each 1mL sample (to make a final concentration of ~1%) and inverted immediately two to three times to ensure complete mixing
3. The samples were incubated for 10 minutes at room temperature with agitation at 8 minutes and then samples were put on a rotating wheel at 40 rpm for the remaining 2 minutes
4. Crosslinking was stopped by adding 115µL glycine (Diagenode, cat no.: C01010130) and incubating for 5 minutes at room temperature
5. The samples were centrifuged at 300 x g at 4°C for 10 minutes using a low deceleration speed
6. The supernatant was aspirated leaving approximately 30µL of the solution and pellet
7. 5472.5µL Hanks' balanced salt solution (HBSS) (Fisher Scientific cat. no.: 11570616) and 27.5µL PIC were added together. Cells were washed with 1mL of ice cold HBSS plus PIC and inverted four to five times to re-suspend cells
8. The samples were centrifuged at 300 x g for 10 minutes at 4°C. The supernatant was discarded and the cell pellet was kept on ice

2.10.1.3. Cell lysis and chromatin shearing

1. Lysis buffer was prepared by adding 522.38µL lysis buffer (Diagenode, cat no.: C01010130) and 2.63µL PIC together. 100µL of complete lysis buffer was added per sample. The samples were inverted to re-suspend the cells and allow bubbles to form. Before being incubated on ice for 10 minutes to ensure complete cell lysis
2. 1641.75µL HBSS and 8.25µL PIC were combined and 300µL of the resultant complete HBSS was added to each sample and inverted four to five times.
3. The samples were then split into 200µL aliquots
4. Half of the aliquots were left on ice, while the remaining half were sonicated for 35 minutes (cycles of 30 seconds on, 30 seconds off, on high power), making sure the water temperature was kept between 4-8°C by using a water cooler. The samples were placed on ice when finished

5. The sonication was repeated with samples that had been left on ice
6. The two 200µL aliquots for each sample were combined and centrifuged for 10 minutes at 14,000 x g at 4°C. The supernatant was then transferred to a new tube

2.10.1.4. Magnetic immunoprecipitation

1. 100µL sheared chromatin was added to the seventh row of the automated IPStar system (Diagenode, Belgium) for each IP
2. 2089.5µL ChIP buffer (Diagenode, cat no.: C01010130) was combined with 10.5µL PIC. 100µL of the resultant complete ChIP buffer was added to each sample in the seventh row
3. 1µg of specific antibody was added to each tube in the seventh row, (equivalent to 1µL H3K4me3 (Diagenode, cat no.: C01010130), 1µL IgG (Diagenode, cat no.: C01010130) or 0.55µL H3K27me3 (Diagenode, cat no.: C15410195)
4. 10µL complete ChIP buffer was also added to each 10µL input sample, which were then stored at 4°C overnight
5. The IPStar was then set up according to an on-screen setup procedure (indirect method). All reagents were provided in the True MicroChIP Kit (Diagenode, cat no.: C01010130)

   Reagent volumes:

   Position A. Elution Buffer 2mL
   Position B. Bead Wash 11mL
   Position C-F. Wash buffer 1-4 3250µL
   Row 3. Magnetic beads 10µL

   Immunoprecipitation conditions:

   1. IP reaction 15 hours at 4°C
   2. Bead Incubation 2 hours at 4°C
   3. Washes 4 minutes at 4°C

   The IPStar was run over night for approximately 21 hours.

2.10.1.5. Elution, decrosslinking and DNA isolation

1. The following morning the samples were recovered from row 12 and placed on the DiaMag02 magnetic rack (Diagenode, Belgium), the supernatant was kept and the beads discarded
2. 100µL of elution buffer (TE1) (Diagenode, cat no.: C01010130) was added to each sample; 180µL was also added to each 20µL input sample from the day before
3. 8µL of elution buffer (TE2) (Diagenode, cat no.: C01010130) was added to each sample
4. All samples were incubated for 4 hours at 65°C, manually agitating every 30 minutes
5. 1000µL of ChIP DNA binding buffer (Diagenode, cat no.: C01010130) was added to each sample and mixed by inversion
6. 625µL of solution was transferred to a spin column in a collection tube (Diagenode, cat no.: C01010130)
7. Samples were centrifuged at 14,000 x g for 1 minute and the flow through was discarded
8. Steps 6 and 7 were repeated with any remaining solution
9. 200µL DNA wash buffer (Diagenode, cat no.: C01010130) was added to each column and centrifuged at 14,000 x g for 1 minute
10. The wash step was then repeated
11. The column was transferred to a new 1.5mL microcentrifuge tube. 40µL of DNA elution buffer (Diagenode, cat no.: C01010130) was applied directly to the column matrix, and incubated at room temperature for 5 minutes. Samples were centrifuged at 14,000 x g for 1 minute to elute the DNA
12. The elution step was repeated so the samples were eluted in a total volume of 80µL
13. Samples were frozen at -20°C ready for qPCR analysis

2.11. Quantitative PCR (qPCR)

qPCR was used to detect a specific DNA target sequence in *ANK1.* This enabled the calculation of the enrichment of a sample relative to the unmodified input standard. Enrichment of specific DNA sequences represented regions in the *ANK1* gene that the histone modification of interest is associated with. Each sample, including input, negative control and standards were run in triplicate for all primers.

2.11.1. qPCR setup

1. 2µL EvaGreen master mix (SolisBiodyne, cat no.: 08-24-00008), 2µL ChIP treated DNA, 1µL forward and reverse primer sequences (Table 5.2) and 5µL of H_2O was combined per sample replicate in 384 well plates (Table 2.20). Seven 1 in 4 serial dilutions and a NTC were included in triplicate per plate

Reagent	Function	Reagent Concentration	Volume (µL)
EvaGreen polymerase	Complete master mix containing magnesium, neucleotides and taq polymerase	5X	2
PCR primers	Short, single-stranded oligonucleotides complementary to target sequence	2/3µM	1
ChIP DNA	ChIP treated DNA	unknown	2
Ultra-pure Water	Ensures accurate and consistent reaction volume	Water to a final volume of 10µL	5

Table 2.20; ANK1 quantitative polymerase chain reaction (qPCR) reagents and volumes.

2.11.2. QuantStudio setup

1. The following qPCR programme was used: 95°C for 10 minutes, 45 cycles of: 95°C for 30 seconds, 60°C for 30 seconds, 72°C for 30 seconds, followed by a dissociation stage

2.11.3. Data normalisation and quality control

1. Standard curves generated by dilution series were used to calculate primer efficiency using the equation:

$$\text{efficiency} = \log \text{dilution}^{(-1/\text{slope})}$$

2. Where slope is the gradient of the correlation between average CT score per replicate and log concentration
3. ChIP data was normalised by calculating the percentage of input. This was done using the equation:

$$\%\ \text{input} = 100 \times \text{primer efficiency}^{((\text{Ct input}-3.32) - \text{Ct IP}))}$$

CHAPTER 3: DNA METHYLOMIC AND HYDROXYMETHYLOMIC SIGNATURES IN ALZHEIMER'S DISEASE BRAIN

The work presented in this chapter is based on work submitted for publication and is currently under review at Genome Biology (Smith et al. 2018). A copy of the submitted manuscript can be found in Appendix 3.

3.1 Introduction

A number of different DNA modifications have been described however; by far the most widely studied is 5mC. This is not surprising considering the abundant, stable and functionally relevant nature of this modification (Bird 2007). It is widely accepted to be involved in a vital capacity in the regulation of genomic expression across human development, ageing and disease. Coupled with the relative ease of profiling this modification using high throughput methods, such as microarrays and sequencing-based technologies, has resulted in the publication of a number of EWAS in various complex diseases. Further research has now shown that other cytosine modifications not only exist but may also have a function in gene regulation (Inoue et al. 2011, Ito et al. 2011). These modifications, including 5hmC, 5fC and 5caC, occur as stages in the de-methylation of cytosine (Figure 1.5).

3.1.1 DNA hydroxymethylation

5hmC was first discovered in bacteriophages in the 1950s (Wyatt et al. 1952). Initially it was believed to be a transient modification, representing an intermediate in the demethylation of 5mC to un-modified cytosine, however recent data suggests that it may have a functional role, with it now having been studied in multiple human tissues (Li et al. 2011, Godderis et al. 2015, Lunnon et al. 2016). Evidence suggests that 5hmC may be particularly important in the brain. The highest levels of 5hmC in any tissue are found in the adult brain (Szwagierczak et al. 2010), whilst a number of studies have shown highly dynamic levels of 5hmC during neurodevelopment (Wang et al. 2012, Spiers et al. 2017). A liquid chromatography mass spectrometry (LC-MS) study that measured 5hmC in mouse brain showed varying levels across the brain, with highest levels in the hypothalamus, cerebral cortex and hippocampus (Munzel et al. 2010). In addition, Kriaucionis et al. (2009) showed 5hmC to be enriched in Purkinje neurones, whilst Khare et al. (2012) showed the modification to be

abundant in synaptic genes. Taken together, it would be interesting to profile 5hmC in the context of neurological disease.

Now that it is becoming increasingly accepted that 5hmC is an independent epigenetic mark, one important question to address relates to the functional consequence of this mark, and its relationship with DNA expression. Recent studies have so far suggested that, like 5mC, it is the location of 5hmC in the genome that is associated with alterations in gene expression. 5hmC has been shown to be enriched in promoters, gene bodies and intergenic areas near genes which in turn positively correlate with gene expression at these loci (Pastor et al. 2011, Song et al. 2011, Wu et al. 2011, Xu et al. 2011). At the chromatin level 5hmC appears to be mainly euchromatic, again suggesting an association with gene transcription (Szulwach et al. 2011). However, this interaction between 5hmC localisation and gene expression is far from conclusive, with some conflicting reports in the literature (Xu et al. 2011).

3.1.2 Role of 5hmC

A growing body of literature indicates that 5hmC is not only an intermediate in the TET catalysed process of active 5mC demethylation, but an epigenetic signal in its own right and contributes to gene regulation (Branco et al. 2012). 5hmC has been shown to be maintained across cellular divisions (Hashimoto et al. 2012) and a number of proteins have been shown to have an equal affinity to 5hmC and 5mC (Frauer et al. 2011), as well as 5hmC specifically (Spruijt et al. 2013). A chaperone protein, ubiquitin like with PHD and ring finger domains 1 (UHRF1), has been found to bind both 5mC and 5hmC with similar affinities and is theorised to promote DNMT1 activity at hydroxymethylated CpGs. In addition, a number of DNA glycosylases and DNA repair proteins have been found to bind to 5hmC, 5fC, and 5caC, however few of these were shown to bind 5mC. Examples of other 5hmC readers have also been documented, including transcriptional regulators, chromatin modifiers and DNA damage repair proteins (Yildirim et al. 2011, Iurlaro et al. 2013, Spruijt et al. 2013); further cementing its role as a functional epigenetic mark. A number of studies suggest that 5hmC has a function in the growth and development of neurones, having being found at relatively high levels in the developing foetal

brain (Lister et al. 2013), as well as being highly abundant in adult brain purkinje neurones of the CER and enriched in synapse-related genes in both mouse and humans (Khare et al. 2012). In addition, a recent study in mouse has shown 5hmC (and 5mC) changes with age (Lardenoije et al. 2015).

3.1.3 5hmC in neurological disease

Changes in the global levels of 5mC and 5hmC have been associated with a number of neurological diseases. To date four publications have reported changes in the global levels of 5mC, 5hmC or both modifications in AD using immunocytochemistry. These results are somewhat contradictory, with numerous studies reporting a global decrease in 5mC, 5hmC or both modifications (Chouliaras et al. 2013, Condliffe et al. 2014, Coppieters et al. 2014); while Lashley et al. (2015) report no detectable changes in AD. Bradley-Whitman et al. (2013) however showed an increase in both 5mC and 5hmC in AD hippocampus. To date only one published study has used DNA sequencing to analyse 5hmC in AD (Zhao et al. 2017). This study utilised a genome wide capture and high-throughput sequencing approach in DNA samples extracted from 30 post-mortem dorsolateral prefrontal cortex tissue samples, identifying 517 differentially hydroxymethylated regions (DHRs) significantly associated with neuritic plaques and 60 DHRs associated with neurofibrillary tangles. However, one caveat of this study is that due to the low resolution of the sequencing analysis, this study was unable to differentiate 5hmC from 5mC. This, coupled with low sample numbers, suggests that these results require further validation. Numerous other neurological diseases have also been linked to changes in 5hmC levels including Autism (Wang et al. 2012, Zhubi et al. 2014), Rett Syndrome (Szulwach et al. 2011), Fragile X Syndrome (Wang et al. 2012) and HD (Villar-Menendez et al. 2013, Wang et al. 2013).

3.1.4 Quantification of 5hmC

To date, it has been difficult to specifically analyse 5hmC levels, in the absence of 5mC levels at base-pair level resolution. Coupling 5hmC antibody immunoprecipitation with sequencing (hMeDIP-seq) is one method that can be used to quantitatively measure 5hmC however, work by Matarese et al. (2011) suggests

that this technique is not yet technically mature and sufficiently standardised to be used for a 5hmC EWAS. Third generation sequencing platforms are another method that could be used as it allows the distinction of modified cytosines in the genome by measuring the change in polymerase kinetics during the incorporation of fluorescently labelled nucleotides (Flusberg et al. 2010). However, these methods are still under development, costly and time-consuming.

Two recent EWAS have utilised the 450K array to demonstrate robust and reproducible changes in 5mC at a number of loci in the AD brain (De Jager et al. 2014, Lunnon et al. 2014), including the *ANK1* gene. However, one caveat of this approach is the utilisation of BS-treated DNA, which converts unmodified cytosines to thymine, whilst methylated cytosines are unchanged. Although the focus of these studies has been on DNA methylomic dysfunction, 5hmC is also protected from conversion, and as such these published EWAS actually report a sum of 5mC and 5hmC levels (Huang et al. 2010).

Recent studies have demonstrated that oxidative-BS (OxBS) treatment enables the conversion of 5hmC to thymine, and by running matched BS and OxBS treated samples in parallel one can generate a quantitative measurement for total DNA modifications (BS data), 5mC (OxBS data) and, by proxy, 5hmC (BS data – OxBS data) (Booth et al. 2012, Booth et al. 2013). At least two publications have now utilised this approach in combination with the 450K array platform to demonstrate brain region-specific differences in DNA modifications in post-mortem tissue from non-demented individuals (Field et al. 2015, Lunnon et al. 2016).

3.1.5 Summary

Although EWAS in AD have reported alterations in 5mC at specific loci, these studies have utilised BS-treated DNA and, as such, actually report a sum of both 5mC and 5hmC. Immunocytochemistry has shown global levels of 5hmC are altered in AD brain, with 5hmC present in neurones, and enriched in genes with synapse-related functions. However, to date no study has examined the levels of 5hmC in AD at single base resolution due to appropriate genetic tools being unavailable or too costly for sequencing an entire human genome. As 5hmC has been shown to be

enriched in the human brain, particularly in neurones and synaptic genes, the study of 5hmC in AD may lead to a better understanding of the molecular mechanisms that underlie this disease.

3.2 Aims

- To perform an EWAS of 5mC and 5hmC levels in parallel in AD EC and CER.
- To assess global 5hmC and 5mC levels in AD EC and CER.
- To identify differentially modified positions (DMoPs), DMPs and differentially hydroxymethylated positions (DHPs) associated with AD neuropathology (Braak stage).
- To compare Braak-associated DMoPs identified in the EC and CER data with previously published EWAS EC and CER data in an independent study cohort (Lunnon et al. 2014) to check for reproducibility.
- To identify DMRs and differentially hydroxymethylated regions (DHRs), consisting of multiple Braak-associated DMPs and DHPs, respectively.
- To validate the most interesting loci (*ANK1*) using another technology (pyrosequencing (Qiagen)) in an independent cohort.

3.3 Materials and Methods

3.3.1 Subjects and samples

For BS and OxBS Illumina 450K profiling (discovery cohort) we used matched brain samples collected from the EC and CER from 96 individuals archived in LNDBB. These samples were not utilised in previously published AD EWAS publications (Lunnon et al. 2014, Smith et al. 2016, Smith et al. 2018). Our first validation cohort consisted of previously published (Lunnon et al. 2014) BS EWAS data generated in matched EC (*n* = 104) and CER (*n* = 108) brain tissue. Our second validation cohort (for pyrosequencing) consisted of matched EC samples from 96 individuals archived in the Oxford Brain Bank Collection (Esiri 1993). For all cohorts, individuals had varying degrees of AD pathology (Braak Stage 0-VI) and all AD patients were over the age of 65 at the time of clinical diagnosis. All samples were dissected by trained specialists, snap-frozen and stored at -80°C. Further information about all samples is provided in Table 3.1. For the discovery cohort and validation cohort 2, genomic DNA was isolated from ~100mg of each dissected brain region using a standard phenol-chloroform extraction method as described in Section 2.3, and tested for degradation and purity prior to analysis as described in Section 2.7.

A)

		Entorhinal Cortex (EC)			Cerebellum (CER)		
		Total DNA Modifications (BS)	5mC (OxBS)	5hmC (BS - OxBS)	Total DNA Modifications (BS)	5mC (OxBS)	5hmC (BS - OxBS)
N passed QC		91	85	85	95	94	94
Sex [M/F] (%)		(56) 51/40 (44)	(56.5) 48/37 (43.5)	(56.5) 48/37 (43.5)	(56.8) 54/41 (43.2)	(57.4) 54/40 (42.6)	(57.4) 54/40 (42.6)
Mean age (±SD)		81.2 (9.5)	81.3 (9.5)	81.2 (9.5)	81.2 (9.3)	81.2 (9.3)	81.2 (9.3)
Diagnosis [AD/CTL] (%)		(70.3) 64/27 (29.7)	(70.6) 60/25 (29.4)	(70.6) 60/25 (29.4)	(70.5) 67/28 (29.5)	(70.2) 66/28 (29.8)	(70.2) 66/28 (29.8)
Braak Stage	0	8	7	7	8	8	8
	I	3	3	3	3	3	3
	II	11	10	10	12	12	12
	III	6	6	6	6	6	6
	IV	8	7	7	10	10	10
	V	18	17	17	19	18	18
	VI	37	35	35	37	37	37
Mean postmortem Interval [minutes] (±SD)		2539.5 (1288.1)	2490.7 (1288.5)	2490.7 (1288.5)	2576.5 (1315.2)	2581.6 (1321.3)	2581.6 (1321.3)

Table 3.1

B)

		Entorhinal Cortex (EC)	Cerebellum (CER)
		Total DNA Modifications (BS)	
N passed QC		104	108
Sex [M/F] (%)		(40.4) 42/62 (59.6)	(42.6) 46/62 (57.4)
Mean age (±SD)		84.9 (8.7)	83.8 (10.0)
Diagnosis [AD/CTL] (%)		(59.6) 62/42 (40.4)	(59.3) 64/44 (40.7)
Braak Stage	0	5	8
	I	11	11
	II	8	7
	III	13	13
	IV	5	5
	V	18	19
	VI	44	45
Mean postmortem Interval [minutes] (±SD)		1997.6 (1227.0)	2071.7 (1303.6)

Table 3.1 cont.

C)

		Entorhinal Cortex (EC)		
		Total DNA Modifications (BS)	5mC (OxBS)	5hmC (BS - OxBS)
N passed QC		96	92	92
Sex[M/F] (%)		(56.3) 54/42 (43.8)	(56.5) 52/40 (43.5)	(56.5) 52/40 (43.5)
Mean age (±SD)		85.0 (7.2)	84.8 (7.3)	84.8 (7.3)
Diagnosis [AD/CTL] (%)		(50) 48/48 (50)	(50) 46/46 (50)	(50) 46/46 (50)
Braak Stage	0	5	5	5
	I	6	6	6
	II	37	35	35
	III	0	0	0
	IV	0	0	0
	V	23	22	22
	VI	25	24	24
Mean postmortem Interval [minutes] (±SD)		2960.6 (1943.8)	2958.9 (1958.7)	2958.9 (1958.7)

Table 3.1: Sample and demographic information for the three cohorts used. *(A) The discovery cohort consisted of 450K array BS and OxBS data generated in 96 individuals from the MRC London Brain Bank for Neurodegenerative Disease. (B) Validation cohort 1 consisted of previously published 450K array BS data generated in an independent cohort of 117 individuals also from the MRC London Brain Bank for Neurodegenerative Disease. (C) Validation cohort 2 consisted of pyrosequencing BS and OxBS data we generated in an independent cohort of 96 individuals from the Thomas Willis Oxford Brain Collection. For the discovery cohort and validation cohort 1 we analysed data from two brain regions: the EC and CER. For validation cohort 2 we analysed ANK1 pyrosequencing data from the EC only. Shown for each dataset are the number of samples (N) that passed Quality control (QC), distribution of sex, clinical diagnosis (AD) or non-demented control (CTL)), Braak stage, Mean age and Post-mortem interval (± standard deviation (SD)).*

3.3.2 Methylomic and hydroxymethylomic profiling

1µg DNA from each sample (discovery cohort and validation cohort 2) was treated with sodium BS and OxBS in parallel using the true-methyl CEGX 96 kit (CEGX, Cambridge, UK) according to the manufacturer's standard protocol, as described in Section 2.1.3. Briefly, DNA samples were split, with half being oxidised (OxBS) and the remainder (BS) going through a mock oxidisation step, before all being BS treated. Gel electrophoresis of the DNA after a digestion control step (Section 2.5.3) was used to test for conversion efficiency (Figure 3.1). All samples were then processed using the 450K array (Illumina Inc, CA, USA) according to the manufacturer's instructions, with minor amendments and quantified using an Illumina HiScan System (Illumina, CA, USA), as described in Section 2.8. Corresponding OxBS and BS treated DNA for the same sample were run together on the same BeadChip. Samples were assigned a unique code for the purpose of the experiment and grouped by tissue. Samples were randomised in their OxBS and BS pairs with respect to sex and disease status to avoid batch effects, and processed in batches of 12 BeadChips. Illumina Genome Studio software was used to extract the raw signal intensities of each probe (without background correction or normalisation).

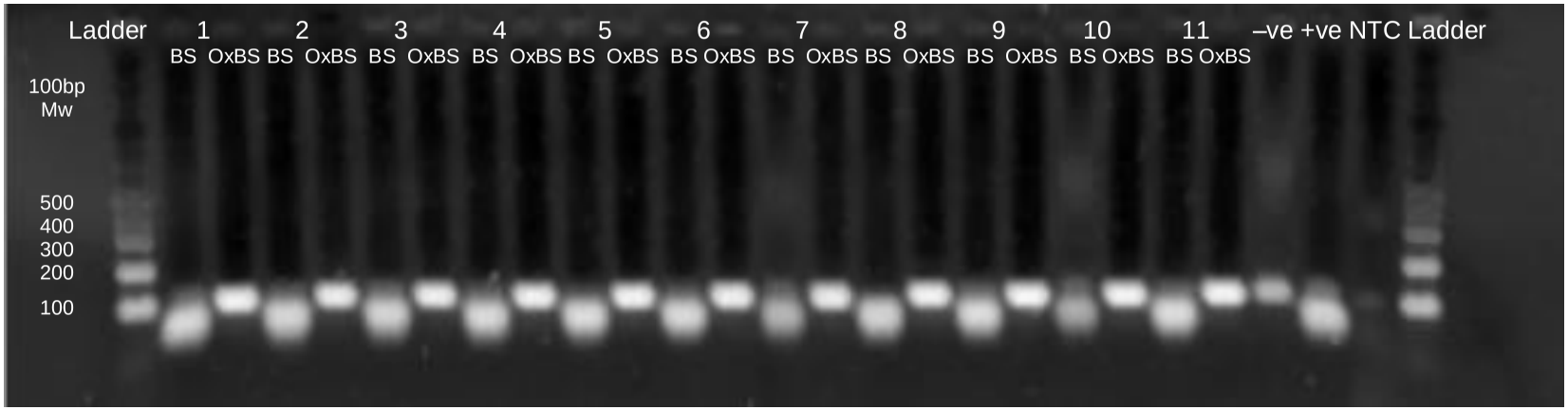


Figure 3.1; Digestion control amplicon conversion gel. *An example gel image of a test amplicon for samples 1-11 after oxidation and digestion control, negative control amplicon (-ve), positive control amplicon (+ve) and no template control (NTC). OxBS treated samples have not been digested by the restriction enzyme and therefore have a higher molecular weight than their BS pair, which have been cleaved by the enzyme and therefore have a lower molecular weight and appear lower on the gel.*

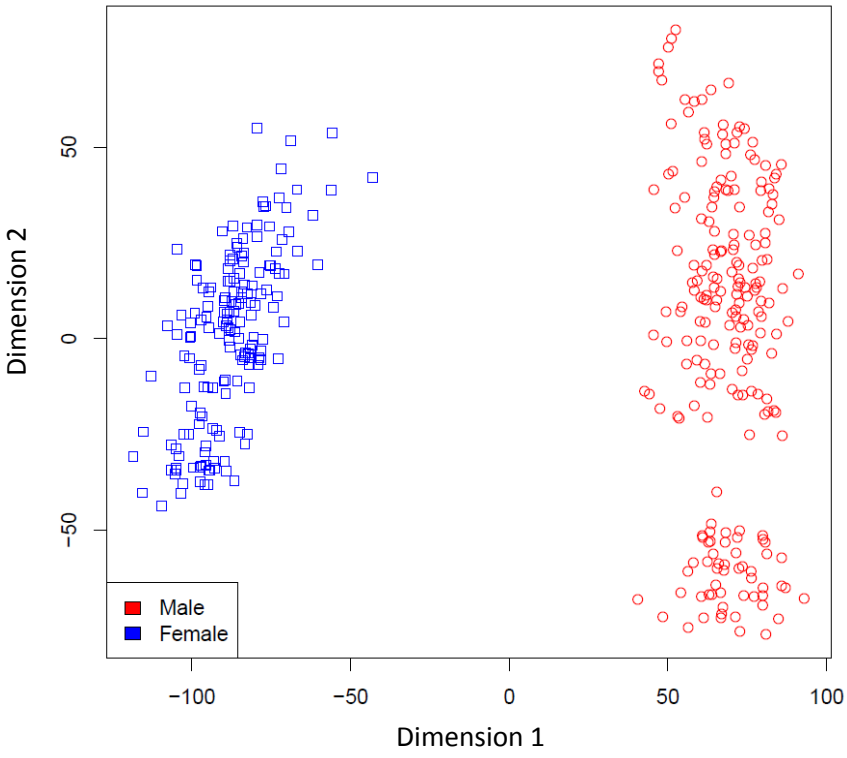
3.3.3 Data quality control

All computations and statistical analyses were performed using R 3.3.2 (R Development Core Team 2012) and Bioconductor 3.5 (Gentleman et al. 2004). Signal intensities were imported into R using the methylumi package (Davis et al. 2012) as a methylumi object. Initial QC checks were performed using functions in the methylumi package.

3.3.3.1 Multi-dimensional scaling

Multi-dimensional scaling of both X and Y chromosomes were used to assess concordance between sex as recorded on post-mortem records and predicted sex using the 65 control SNPs (Figure 3.2). The 65 SNPs were also used to confirm that matched BS EC, OxBS EC, BS CER and OxBS CER samples were sourced from the same individual.

A)



B)

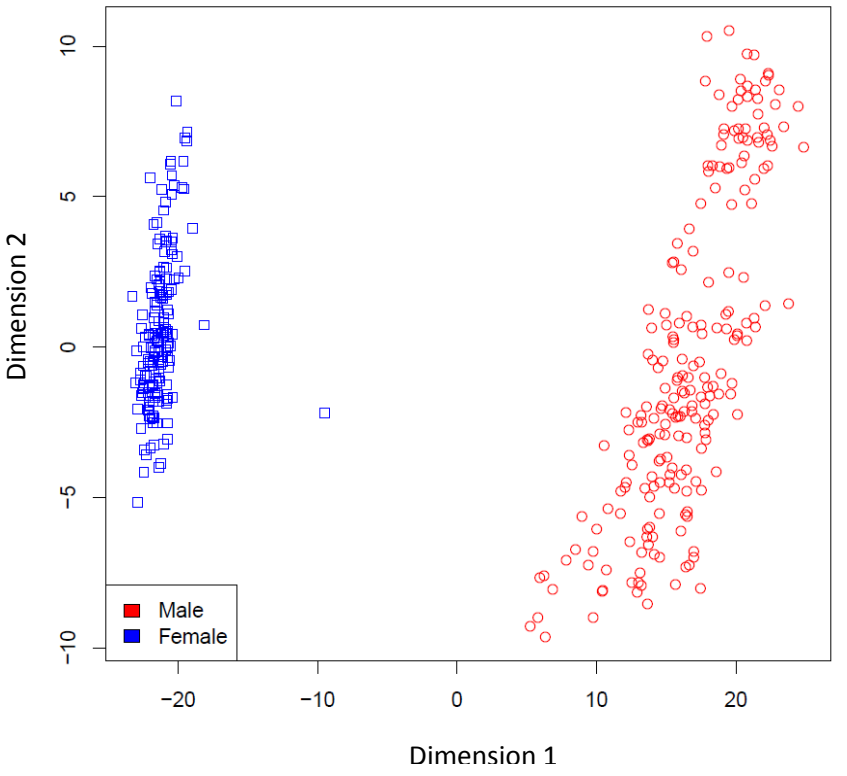


Figure 3.2; Multi-dimensional scaling of sex chromosomes. *(A) X Chromosome scaling, (B) Y Chromosome scaling. This was used to assess concordance between recorded and predicted sex in all samples.*

3.3.3.2 Sample and probe removal with *pfilter*

The *pfilter* function was used to remove samples where 5% of sites had a detection p-value > 0.05. Specific sites were also removed if the beadcount was less than three in 5% of samples or if 1% of the samples had a detection p-value > 0.05 at that position. 15 samples and 38,737 sites were removed at this stage of the analysis.

3.3.3.3 Removal of non-specific and polymorphic probes

Probes with common (minor allele frequency (MAF) > 5%) SNPs in the CG or single base extension position, or probes that are nonspecific or miss-mapped were flagged (N=43,233 probes) and discarded from our results (Chen et al. 2013).

3.3.3.4 Methylation and hydroxymethylation beta densities

Data was pre-processed in the R package *wateRmelon* using the *dasen* function as previously described (Pidsley et al. 2013). Plotting the mean normalised beta density for all BS and OxBS samples revealed a bimodal distribution for both treatments (Figure 3.3). The beta value peak for the OxBS treatment was skewed to the left of the BS converted samples, corresponding to the higher global total modification levels expected for the BS treatment, which included signal from both 5hmC and 5mC. A CER beta value shift to the right was observed in relation to the EC, this was to be expected due to previous reports of high levels of DNA cytosine modifications present in the CER (Kriaucionis et al. 2009).

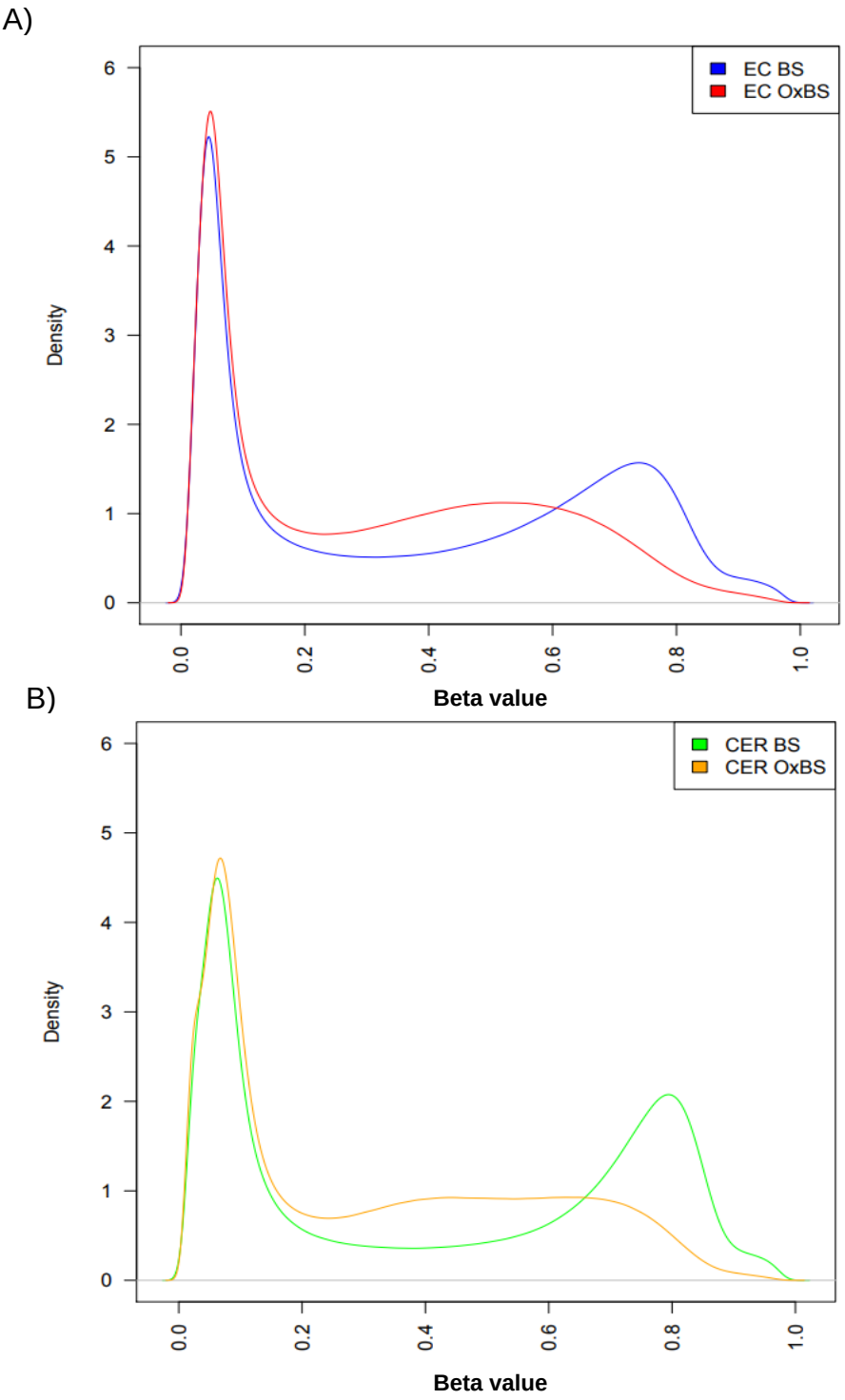


Figure 3.3; Comparison of BS and OxBS data in the EC and CER. *Following data normalisation with the dasen function in the R package WateRmelon we compared the mean betas between BS, and OxBS treated EC (A) and CER (B) samples.*

3.3.4 Data analysis

After these QC steps 367,480 probes were taken forward for analysis. Array data for each tissue was normalised separately and the analysis was performed separately by tissue. To calculate 5hmC levels in a sample, the OxBS signal was subtracted from the BS signal. It has been previously reported that a small proportion of probes in each sample are characterised by a negative BS-oxBS value, likely resulting from technical variance inherent in the Illumina array protocol (Lunnon et al. 2016) and could therefore represent noise in the system. However, we make the assumption that any technical variation present would equally affect both the BS and OxBS signal. Therefore probes with a mean negative 5hmC value across all samples may still represent biological differences and were included in our analysis. We have annotated tables to denote where the mean 5hmC level was >0 (5hmC Positive probe column).

3.3.4.1 Linear regression analysis

The effects of age and sex were regressed out before subsequent analysis. For the identification of DMoPs, DMPs and DHPs specifically altered with respect to neuropathological measures of AD, we performed a quantitative analysis where samples were analysed with respect to Braak stage using a linear regression model. Analysis was repeated to include the bioinformatically derived cell epigenotype specific (CETS) score as a covariate. This score is calculated using the R package “CETS”, which estimates the cell type composition of biological sample based on bulk methylation data (Guintivano et al. 2013). Data was analysed separately for each brain region using linear regression and probes were ranked by *P* value, and Q-Q plots assessed to check for *P* value inflation (Figure 3.4).

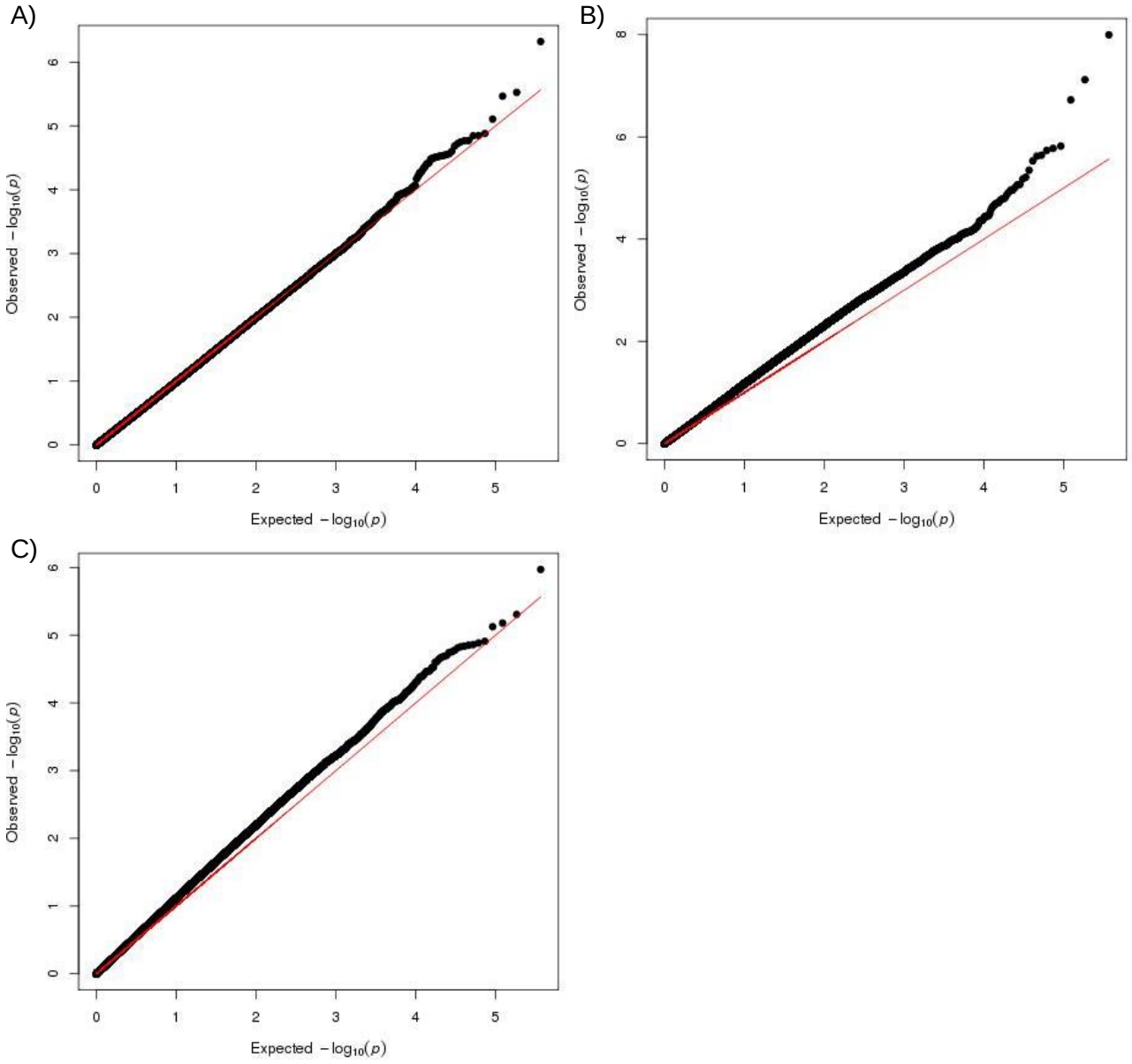


Figure 3.4; Quantile-quantile (Q-Q) plots of expected versus observed P value to check for inflation in linear regression analyses. *The Q-Q plots for (A) total DNA modifications (BS data), (B) 5mC (OxBS data) and (C) 5hmC (BS – OxBS data) in the EC shows no P value inflation in our analyses. Similar Q-Q plots were observed for the CER (data not shown).*

3.3.4.2 Identifying enrichment in specific gene features

To investigate whether our neuropathology-associated DMPs were enriched at specific gene features, we used a Fisher's exact test to test for an enrichment of either nominally-significant hypermethylated or hypomethylated loci ($P < 0.05$) compared to all neuropathology-associated DMPs ($P < 0.05$) at each gene feature as described by Slieker et al (Slieker et al. 2013). Similarly, to investigate whether our neuropathology-associated DHPs were enriched at specific gene features, we used a Fisher's exact test to test for an enrichment of either nominally significant hyperhydroxymethylated or hypohydroxymethylated loci ($P < 0.05$) compared to all neuropathology-associated DHPs ($P < 0.05$) at each gene feature. As this is exploratory work we considered all DMPs and DHPs with $P < 0.05$ to ensure we captured all biological differences.

3.3.4.3 Differentially modified regions

To identify DMRs and DHRs we identified spatially correlated p-values within our data using the Python module *comb-p* (Pedersen et al. 2012) to group ≥ 2 spatially correlated CpGs within a 500bp sliding window with a significance threshold of $P = 0.01$ from our regression analysis. A P value of < 0.01 was used to maintain consistency with previous publication (Lunnon et al. 2014) and to ensure a more stringent identification of DMRs and DHRs.

3.3.4.4 Pathway analysis

Pathway analysis of DMPs and DHPs reaching our significance threshold ($P < 5 \times 10^{-5}$) was performed using the *missmethyl 1.10.0* (Phipson et al. 2016) package from Bioconductor. This P value threshold was used to limit the number of associated genes included, in an attempt to prevent the identification of false positive pathways. Pathways were included if they met the following criteria: (1) the number of genes in the gene list was greater than one and less than 2000, (2) more than one gene from our analysis was present in the pathway and (3) the P value for the association was smaller than 0.05.

3.3.5 Targeted Replication using Bisulfite Pyrosequencing

BS pyrosequencing, as described in Section 2.9, was used to quantify 5mC across eight individual *ANK1* CpG sites, including cg05066959 and cg11823178, spanning from 41519302 to 41519420 within chromosome 8 (hg19). A single amplicon (246bp) was amplified using primers designed using the PyroMark Assay Design software 2.0 (Qiagen, UK) as previously described (Lunnon et al. 2014), and sequenced using two sequencing primers to maximise coverage across eight CpG sites within a 118bp region. 5mC was quantified in validation cohort 2 using the Pyromark Q24 system (Qiagen, UK) following the manufacturer's standard instructions and the Pyro Q24 CpG 2.0.6 software. We used linear regression to assess the association between DNA methylation and low pathology (Braak scores 0-II: n = 48) to high pathology (Braak scores V-VI: n = 48) for total DNA modifications, 5mC and 5hmC at eight individual CpG sites. Linear regression was used to maintain consistency with the data generated on the array and with previous publication (Lunnon et al. 2014). The effects of age and sex were included as covariates in our analysis. For this sample cohort we had no samples with middle stage AD pathology (Braak III-IV). We then averaged across the 118bp amplicon and used two sample t-test to assess the difference in the means between groups. Bonferroni threshold of $P < 0.017$ was used to correct for multiple testing.

3.4 Results

3.4.1 AD entorhinal cortex is characterised by global hypermethylation and hypohydroxymethylation

The majority of studies investigating 5hmC in AD have used immuno-histochemical based methods to quantify global levels (Chouliaras et al. 2013, Condliffe et al. 2014, Coppieters et al. 2014, Lashley et al. 2015, Celarain et al. 2016). Although the 450K array primarily interrogates CG rich regions, which are, in general, unmethylated, we were interested to see if there were any global changes across these regions. To investigate these global levels of DNA modifications in our data we averaged levels across all probes that passed QC (n = 367,480), identifying a global decrease in 5hmC (P = 0.040) and increase in 5mC (P = 0.039), with overall no net change in total modifications (P = 0.838) in the EC with respect to Braak stage. Conversely, no significant Braak stage associated alterations were seen for global levels of 5hmC (P = 0.659), 5mC (P = 0.916) or total modifications (P = 0.366) in the CER (P > 0.05).

3.4.2 Highly reproducible alterations in cytosine modifications are detectable in AD cortex

Initially, we focused on identifying differences in total DNA modification levels (BS data: Table 3.2) in the EC that were associated with Braak stage, whilst controlling for age and sex above our nominated significance threshold ($P < 5 \times 10^{-5}$). To assess the reproducibility of our results, we first compared the effect size of these Braak stage-associated DMoPs from the BS data (Table 3.2), with BS data previously generated on an independent set of samples from the EC in AD (validation cohort 1) (Lunnon et al. 2014). There was a consistent direction of effect between the two studies (Figure 3.5A: sign test P =0.043), with the effect sizes highly correlated between studies (r = 0.71, $P = 4.85 \times 10^{-6}$). Similarly, there was also a significant excess of consistent direction of effect between the two studies when examining the effect size of previously published Braak-associated DMoPs ($P < 5 \times 10^{-5}$) in the EC (Lunnon et al. 2014) with the effect size of the same probes in the current study (Figure 3.5B: sign test $P = 9.77 \times 10^{-04}$), with the effect sizes highly correlated between studies (r = 0.78, $P = 3.50 \times 10^{-4}$).

In contrast, when we compared the effect size of the Braak-associated DMoPs in the CER ($P < 5 \times 10^{-5}$) in the current study (Table 3.3) to BS data previously published in the CER in AD (Lunnon et al. 2014) we found no consistent direction of effect (Figure 3.6A: sign test $P = 0.267$). Similarly, the effect sizes of the previously published CER DMoPs ($P < 5 \times 10^{-5}$) (Lunnon et al. 2014) showed no consistent direction of effect with the same probes in the current study (Figure 3.6B: sign test $P = 0.188$). This suggests that regions of the brain that are relatively protected from neurodegeneration, such as the CER, do not display consistent disease-associated epigenetic changes in total DNA modifications in either a global or loci-specific manner.

Probe information					Total modifications (BS)			5mC (OxBS)		5hmC (BS - OxBS)		
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	*P*	Δ	*P*	Δ	5hmC Positive probe
1	cg21774827	19:54379120	MYADM	PRKCG	**3.46E-07**	7.37	**4.76E-07**	0.066	4.52	0.264	2.68	✓
2	cg26925343	16:28857904	TUFM	SH2B1	**4.06E-06**	-1.28	**2.98E-06**	**0.042**	-1.27	0.998	0.00	·
3	cg17554875	5:179520998	-	RNF130	**4.24E-06**	7.26	**1.77E-05**	0.598	0.83	**8.75E-04**	6.23	✓
4	cg27642528	10:88814651	GLUD1	GLUD1	**5.75E-06**	3.71	**1.89E-05**	0.825	0.28	**0.005**	3.55	✓
5	cg20066612	1:160050948	KCNJ9	KCNJ9	**6.79E-06**	3.25	**2.73E-05**	**0.027**	-2.89	**2.70E-05**	5.84	✓
6	cg03975188	7:148131599	-	CUL1	**7.00E-06**	8.28	**2.07E-05**	0.867	0.33	**0.002**	8.39	✓
7	cg17304222	5:180076905	FLT4	FLT4	**8.29E-06**	-1.16	**1.71E-05**	**0.002**	-1.07	0.860	-0.07	·
8	cg05066959	8:41519308	ANK1	NKX6-3	**8.37E-06**	10.98	**2.94E-05**	**8.03E-05**	13.71	0.349	-2.72	✓
9	cg01042637	3:130743723	ASTE1	ASTE1	**9.75E-06**	6.12	**3.20E-05**	0.172	2.65	0.112	3.39	✓
10	cg04658038	17:64800166	PRKCA	CACNG5	**1.08E-05**	4.48	**1.32E-05**	0.634	0.87	0.068	3.72	✓
11	cg06653632	12:129281444	SLC15A4	SLC15A4	**1.21E-05**	6.87	**4.47E-05**	0.151	3.00	0.083	3.81	✓
12	cg04904331	7:49813033	VWC2	VWC2	**1.59E-05**	-1.15	**2.51E-05**	0.170	0.43	**2.89E-04**	-1.54	·
13	cg21151057	7:157393764	PTPRN2	DNAJB6	**1.59E-05**	5.49	**5.19E-05**	0.983	0.04	**0.027**	5.54	✓
14	cg17867333	5:178423163	GRM6	GRM6	**1.87E-05**	-5.21	**3.08E-05**	0.148	-1.94	**0.007**	-3.30	·
15	cg13851211	16:50321678	ADCY7	ADCY7	**2.91E-05**	7.09	**2.87E-05**	**0.034**	4.13	0.224	2.86	✓
16	cg15197657	5:176916434	PDLIM7	PDLIM7	**3.28E-05**	-6.43	**9.83E-05**	0.183	-5.18	0.820	-0.84	·
17	cg20618448	19:49962324	ALDH16A1	ALDH16A1	**3.29E-05**	5.38	**3.33E-05**	**0.001**	5.94	0.721	-0.44	✓
18	cg06898199	7:128502890	ATP6V1F	ATP6V1F	**3.44E-05**	1.17	**1.69E-05**	**0.017**	0.68	0.154	0.47	·
19	cg23022785	20:10153355	-	SNAP25	**3.68E-05**	-1.50	**6.06E-05**	0.375	-0.52	0.124	-1.02	·
20	cg06192381	3:50305066	SEMA3B	LSMEM2	**4.26E-05**	-1.33	**1.54E-04**	0.316	0.43	**0.001**	-1.62	·
21	cg00627837	22:19133638	DGCR14	DGCR14	**4.32E-05**	-4.18	**1.14E-04**	0.431	-0.94	**0.033**	-3.11	·
22	cg26040816	20:42816135	JPH2	JPH2	**4.62E-05**	-2.49	**1.03E-04**	0.787	0.28	**0.010**	-2.59	·
23	cg01348086	5:176785296	RGS14	RGS14	**4.73E-05**	4.13	**8.67E-05**	**4.87E-04**	3.08	0.163	1.10	✓
24	cg03309308	16:68525329	-	SMPD3	**4.78E-05**	-4.88	**1.13E-04**	**0.009**	-3.08	0.277	-1.61	·
25	cg10794439	8:41518051	ANK1;MIR486	NKX6-3	**4.81E-05**	6.69	**1.98E-04**	**0.022**	6.70	0.901	0.32	✓

Table 3.2: Significant neuropathology-associated differentially modified positions (DMoPs) in the EC. *Shown for each probe are chromosomal location (hg19), UCSC annotation, GREAT annotation, with corrected effect size (difference (Δ) in total DNA modification levels) and corresponding P value between Braak score 0 and Braak score VI after adjusting for the covariates of age and sex, and an additional corresponding P value also correcting for NeuN+ neuronal proportions (P CETS corrected). Shown for all probes are Δ and corresponding P value for 5mC and 5hmC levels alone, with annotation as to whether the mean 5hmC level was >0. All P values <0.05 are shown in bold.*

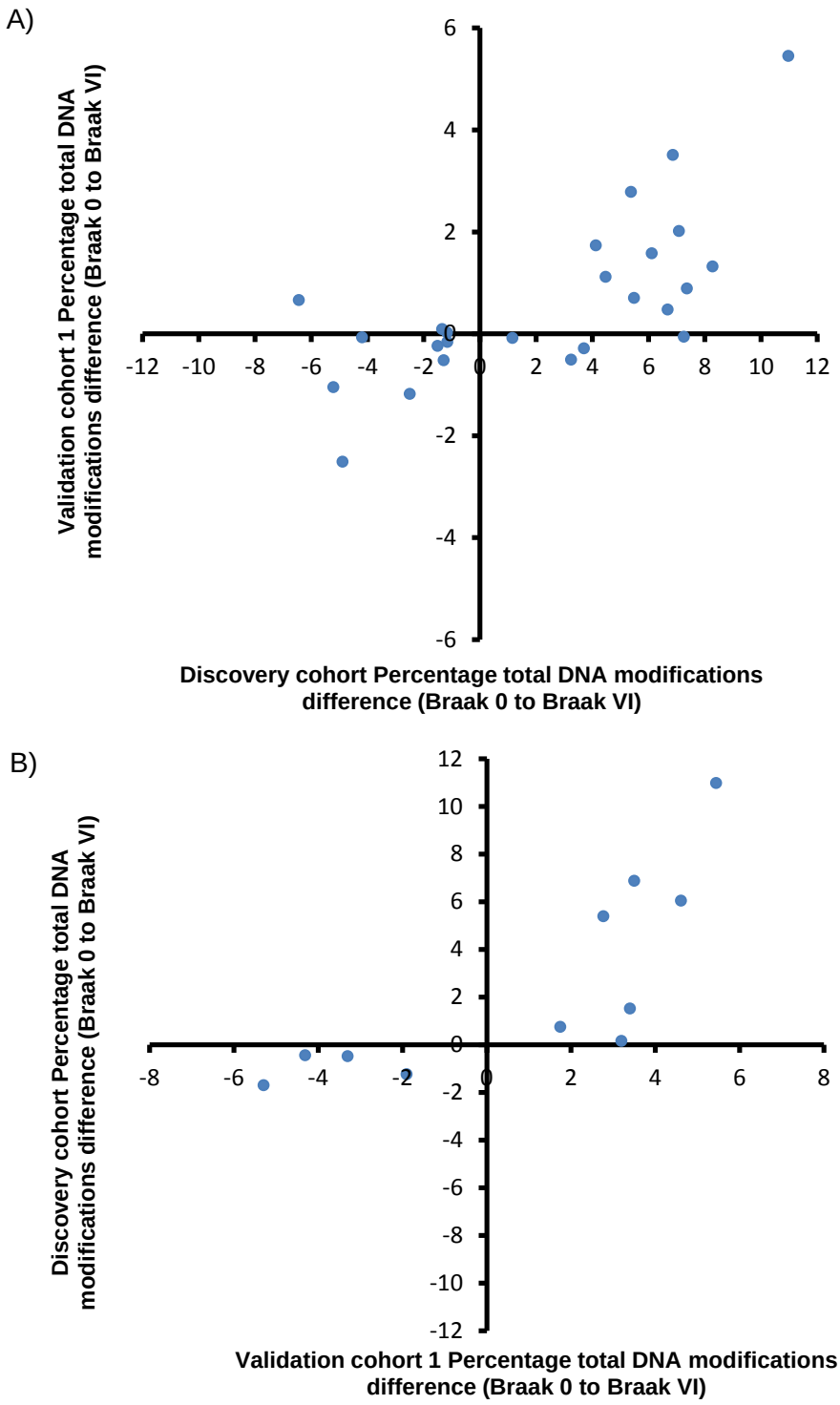


Figure 3.5; Braak-associated EC DMoPs identified in this study are consistent with those identified in previous analyses of AD brain. *(A) There was a consistent direction of effect for Braak-associated DMoPs ($P < 5 \times 10^{-5}$) identified in the EC in the discovery cohort when compared with the same probes in a previously published dataset of Braak-associated DMoPs in the EC (validation cohort 1) (sign test P =0.043), with the effect sizes highly correlated between studies (Pearson's correlation df = 98, r = 0.71, $P = 4.85 \times 10^{-6}$). (B) Similarly, there was a consistent direction of effect for previously published Braak-associated DMoPs in the EC ($P < 5 \times 10^{-5}$) (validation cohort 1) when compared with the same probes in the current study (sign test $P = 9.77 \times 10^{-04}$), with the effect sizes highly correlated between studies (Pearson's correlation df = 98, r = 0.78, $P = 3.50 \times 10^{-4}$).*

Probe information					Total DNA modifications (BS)			5mC (OxBS)		5hmC (BS - OxBS)		
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	*P*	Δ	*P*	Δ	5hmC Positive probe
1	cg26572042	1:21113758	HP1BP3	HP1BP3	**1.45E-06**	-1.00	**6.13E-07**	**0.001**	-0.62	0.153	-0.38	·
2	cg17598713	6:33265534	RGL2	RGL2	**4.83E-06**	-2.25	**2.08E-06**	0.279	-0.78	0.073	-1.48	·
3	cg18107396	6:35436198	RPL10A	RPL10A	**1.02E-05**	0.56	**1.37E-05**	0.142	0.20	**0.050**	0.36	·
4	cg16085178	6:135818824	AHI1;C6orf217	AHI1	**1.49E-05**	0.67	**2.03E-05**	0.798	0.07	**0.049**	0.61	·
5	cg03415001	4:181480012	-	-	**2.55E-05**	-3.42	**2.27E-05**	0.537	0.53	**0.001**	-3.89	✓
6	cg04074793	10:129924374	MKI67	MKI67	**2.57E-05**	-0.72	**3.33E-05**	0.841	-0.04	**0.009**	-0.68	·
7	cg16720675	18:19284817	ABHD3	ABHD3	**3.07E-05**	0.96	**3.50E-05**	0.927	-0.05	**0.041**	1.00	·
8	cg02215070	7:134144055	AKR1B1	AKR1B1	**3.59E-05**	0.79	**3.66E-05**	0.407	0.15	**0.009**	0.64	✓
9	cg04899175	12:117798748	NOS1	NOS1	**3.87E-05**	0.92	**5.12E-05**	**0.006**	1.07	0.708	-0.15	·
10	cg08161720	3:52188073	WDR51A	POC1A	**3.89E-05**	-1.41	**4.18E-05**	0.377	-0.44	0.101	-0.95	·
11	cg14289985	19:57018848	ZNF471	ZNF471	**4.17E-05**	1.07	**5.65E-05**	0.161	-0.44	**2.61E-05**	1.53	✓
12	cg24022821	6:28831886	-	TRIM27	**4.35E-05**	0.89	**5.40E-05**	**0.036**	0.50	0.144	0.40	✓
13	cg07612045	7:144533234	TPK1	TPK1	**4.89E-05**	-0.79	**5.70E-05**	0.899	-0.05	0.126	-0.74	·

Table 3.3; Significant neuropathology-associated differentially modified positions (DMoPs) in the CER. *Shown for each probe are chromosomal location (hg19), UCSC annotation, GREAT annotation, with corrected effect size (difference (Δ) in total DNA modification levels) and corresponding linear regression derived P value between Braak score 0 and Braak score VI after adjusting for the covariates of age and sex, and an additional corresponding P value also correcting for NeuN+ neuronal proportions (P CETS corrected). Shown for all probes are Δ and corresponding P value for 5mC and 5hmC levels alone, with annotation as to whether the mean 5hmC level was >0. All P values <0.05 are shown in bold.*

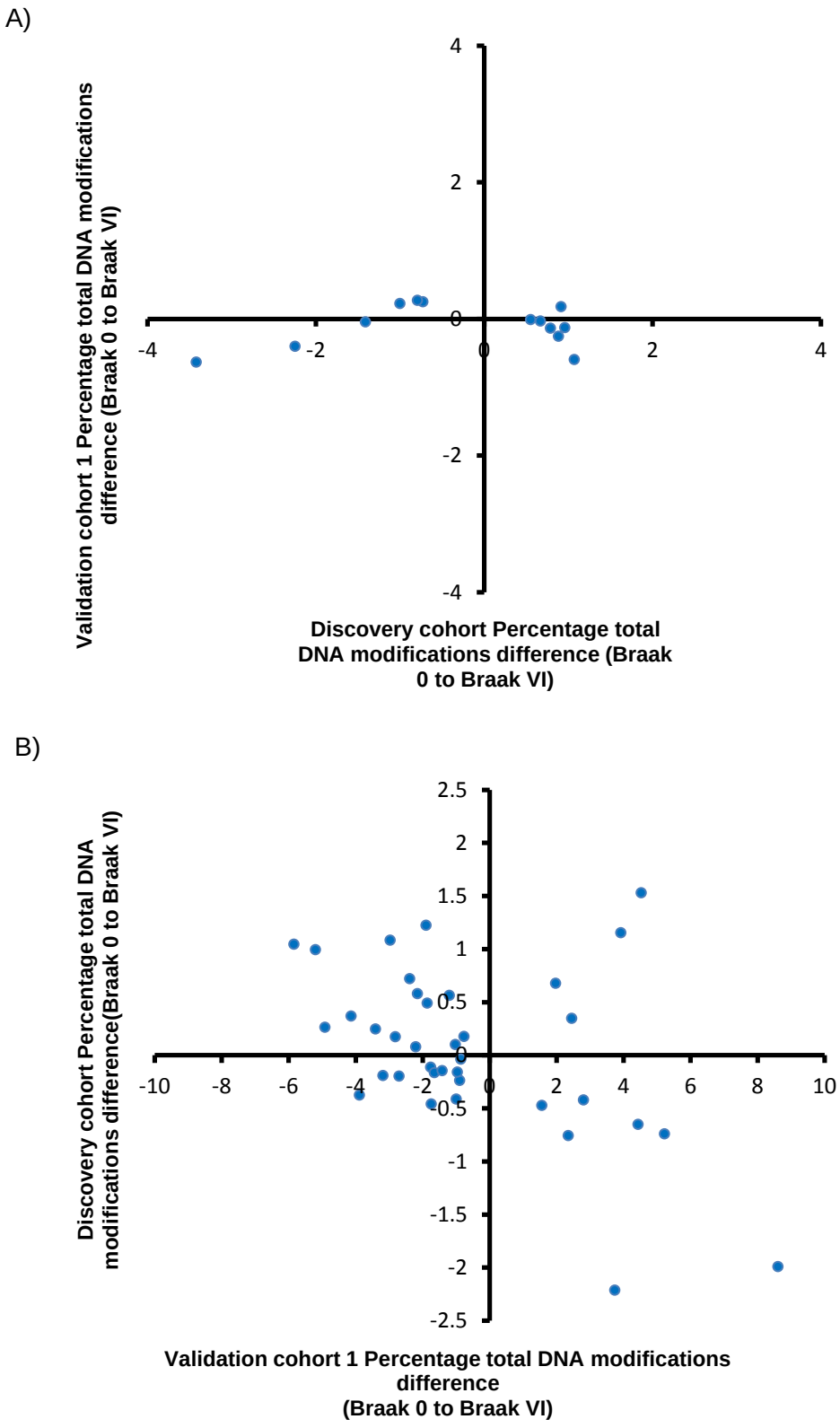


Figure 3.6; Braak-associated DMoPs in the CER are not consistent between studies *(A) There was no consistent direction of effect when comparing Braak-associated DMoPs ($P < 5 \times 10^{-5}$) in the CER in the discovery cohort with the effect size of the same probes in a previously published dataset of Braak-associated DMoPs in the CER (validation cohort 1) (sign test P =0.267). (B) There was also no consistent direction of effect for previously published Braak-associated DMoPs in the CER ($P < 5 \times 10^{-5}$) (validation cohort 1) when compared with the effect size of the same probes in the current study (sign test P = 0.188).*

3.4.3 Locus specific changes in DNA methylation and hydroxymethylation occur in AD

We were interested to identify specific loci that showed Braak-associated differential 5mC and 5hmC in the EC (Tables 3.4-3.5; Figure 3.7). We identified three Braak-associated DMPs that reached experiment-wide significance ($P < 2.2 \times 10^{-7}$), with a further 72 loci reaching our nominal significance threshold of $P < 5 \times 10^{-5}$. Although no Braak-associated differentially hydroxymethylated positions (DHPs) reached experiment wide significance, we did identify 57 loci that reached our nominal significance threshold. Of the three experiment-wide significant DMPs, the first and third most significant were found on chromosome 3 and resided in *CGGBP1* and *GNL3* respectively, whilst the second most significant DMP was located in *WNT5B* on chromosome 12. In the CER we only identified seven Braak-associated DMPs (Table 3.6) and 12 DHPs (Table 3.7) that reached our nominal significance threshold ($P < 5 \times 10^{-5}$).

Probe information					5mC (OxBS)			Total modifications (BS)		5hmC (BS - OxBS)		
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	*P*	Δ	*P*	Δ	5hmC Positive probe
1	cg15147060	3:88108213	CGGBP1	CGGBP1	**3.65E-09**	-1.92	**1.01E-08**	0.440	-0.13	**3.94E-07**	1.73	·
2	cg10696062	12:1726028	WNT5B	WNT5B	**2.82E-08**	13.73	**7.60E-08**	**0.009**	4.45	**1.07E-04**	-9.58	·
3	cg00106685	3:52720133	GNL3;PBRM1	GNL3	**7.80E-08**	-1.45	**1.89E-07**	0.413	-0.13	**4.04E-06**	1.39	·
4	cg18110359	2:45179970	-	SIX3	**5.57E-07**	-2.18	**1.51E-06**	0.150	-0.60	**1.33E-04**	1.95	·
5	cg15624314	1:226251003	H3F3A;LOC440926	H3F3A	**6.14E-07**	-1.84	**1.68E-06**	0.320	-0.25	**4.30E-05**	1.59	·
6	cg17026303	3:38495161	LOC100128640;ACVR2B	ACVR2B	**6.70E-07**	-2.46	**1.84E-06**	0.948	-0.02	**2.74E-06**	2.39	·
7	cg07919466	19:50833956	KCNC3	KCNC3	**8.41E-07**	-1.83	**2.28E-06**	0.730	-0.09	**1.03E-05**	1.69	·
8	cg05052335	5:130970657	RAPGEF6	RAPGEF6	**9.73E-07**	-1.56	**2.38E-06**	0.215	-0.17	**6.00E-06**	1.39	·
9	cg26332552	3:197402549	MIR922;KIAA0226	KIAA0226	**1.08E-06**	11.17	**2.95E-06**	**4.17E-04**	6.57	**0.037**	-4.40	✓
10	cg03503087	10:118032626	GFRA1	GFRA1	**2.38E-06**	-1.89	**4.50E-06**	0.731	0.06	**1.02E-05**	1.97	·
11	cg02456292	19:42817262	TMEM145	TMEM145	**3.26E-06**	-2.72	**6.15E-06**	0.147	-0.49	**7.63E-04**	2.22	·
12	cg04418576	14:103851554	MARK3	MARK3	**3.28E-06**	-1.67	**8.53E-06**	0.218	-0.23	**1.86E-04**	1.41	·
13	cg18070033	2:220025364	NHEJ1	NHEJ1	**3.64E-06**	-1.84	**8.65E-06**	0.919	0.02	**6.24E-06**	1.84	·
14	cg05341384	16:86331661	-	FOXF1	**4.15E-06**	-4.92	**1.10E-05**	0.875	-0.10	**2.01E-05**	4.41	·
15	cg14467668	3:52571040	LOC440957	SMIM4	**4.66E-06**	-1.38	**1.22E-05**	0.874	0.03	**1.94E-04**	1.34	·
16	cg12144476	21:27541894	APP	APP	**4.85E-06**	-1.39	**9.79E-06**	0.659	0.08	**1.61E-05**	1.44	·
17	cg12366118	12:62860687	MON2	MON2	**4.96E-06**	-1.73	**1.35E-05**	0.126	-0.31	**2.07E-04**	1.39	·
18	cg14294321	19:12895288	-	JUNB	**6.43E-06**	7.16	**1.57E-05**	0.274	1.45	**0.005**	-5.58	✓
19	cg00639289	2:213403734	ERBB4	ERBB4	**6.86E-06**	-1.32	**1.62E-05**	0.418	-0.11	**5.12E-04**	1.11	·
20	cg24516830	1:71694460	-	ZRANB2	**6.89E-06**	9.60	**1.66E-05**	0.122	2.57	**0.003**	-6.74	✓
21	cg00847044	4:2419822	ZFYVE28	ZFYVE28	**7.22E-06**	-1.82	**1.93E-05**	0.974	-0.01	**1.66E-04**	1.77	·
22	cg25329939	13:98085947	RAP2A	RAP2A	**7.73E-06**	-1.29	**1.74E-05**	0.074	0.31	**2.05E-06**	1.56	·
23	cg03792042	8:145045235	PLEC1	PARP10	**7.95E-06**	10.13	**2.10E-05**	0.162	3.03	**0.002**	-6.85	✓
24	cg06624337	15:47477411	-	SEMA6D	**9.74E-06**	-1.64	**2.63E-05**	0.460	-0.17	**0.001**	1.41	·
25	cg03263685	2:68480160	PPP3R1	PPP3R1	**1.17E-05**	-1.23	**3.08E-05**	0.932	0.01	**1.28E-05**	1.21	·
26	cg01888395	9:132145105	-	NTMT1	**1.20E-05**	7.14	**2.38E-05**	0.068	3.10	**0.017**	-4.13	✓
27	cg22782712	10:9920178	-	-	**1.27E-05**	5.39	**2.27E-05**	0.477	-0.93	**1.82E-04**	-6.71	·
28	cg05645927	6:27344393	ZNF204P	ZNF391	**1.29E-05**	7.21	**6.68E-06**	0.793	0.40	**0.001**	-6.64	·
29	cg14102251	14:75745098	FOS	FOS	**1.32E-05**	-1.67	**3.51E-05**	0.487	-0.15	**8.36E-05**	1.50	·
30	cg02412050	6:35995429	MAPK1	MAPK14	**1.39E-05**	-1.41	**3.63E-05**	**0.031**	-0.39	**0.007**	0.94	·
31	cg24631913	5:87971905	LOC645323	MEF2C	**1.52E-05**	-1.41	**3.62E-05**	0.598	-0.12	**5.80E-04**	1.25	·
32	cg23648516	13:78467552	-	EDNRB	**1.70E-05**	9.69	**4.27E-05**	0.673	-0.63	**5.31E-05**	-10.48	✓
33	cg21902245	X:79270180	TBX22	TBX22	**1.96E-05**	-6.42	**4.51E-05**	**0.017**	-4.85	0.420	1.29	✓

Table 3.4.

Probe information					5mC (OxBS)			Total modifications (BS)		5hmC (BS - OxBS)		
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	*P*	Δ	*P*	Δ	5hmC Positive probe
34	cg16898576	11:63449499	RTN3	RTN3	**2.09E-05**	-1.31	**5.45E-05**	0.512	-0.11	**7.91E-04**	1.15	·
35	cg15235057	19:16683989	SLC35E1	SLC35E1	**2.31E-05**	-1.46	**3.52E-05**	0.522	-0.12	**2.96E-04**	1.26	·
36	cg03160445	3:100120349	LNP1;TOMM70A	TOMM70A	**2.46E-05**	-1.48	**6.34E-05**	0.680	0.07	**5.35E-06**	1.50	·
37	cg11823178	8:41519399	ANK1;MIR486	NKX6-3	**2.61E-05**	9.31	**6.64E-05**	**4.68E-04**	6.03	0.068	-3.20	✓
38	cg11507178	11:118781515	BCL9L	BCL9L	**2.63E-05**	-1.92	**7.07E-05**	0.219	-0.59	**0.039**	1.22	·
39	cg13876163	1:45140495	C1orf228;TMEM53	TMEM53	**2.73E-05**	-2.25	**4.28E-05**	0.145	0.33	**8.39E-06**	2.55	·
40	cg16933762	3:77147109	ROBO2	-	**2.77E-05**	10.36	**7.41E-05**	0.493	-1.12	**2.80E-05**	-10.61	·
41	cg18282456	14:51561492	TRIM9	TRIM9	**2.85E-05**	-2.13	**7.52E-05**	0.430	-0.23	**7.62E-04**	1.83	·
42	cg11530914	16:67281528	FHOD1;SLC9A5	FHOD1	**2.86E-05**	-1.87	**7.39E-05**	0.642	0.12	**7.74E-06**	1.95	·
43	cg23414228	1:155043413	-	EFNA4	**2.96E-05**	-1.62	**7.36E-05**	0.627	-0.12	**1.11E-04**	1.48	·
44	cg13323756	7:29847108	-	WIPF3	**2.97E-05**	3.90	**1.10E-05**	0.398	0.88	**0.015**	-2.39	✓
45	cg15268622	1:167189846	POU2F1	POU2F1	**3.00E-05**	-1.13	**7.54E-05**	0.247	-0.15	**2.50E-04**	1.00	·
46	cg09811464	9:135465645	-	BARHL1	**3.04E-05**	-2.86	**3.49E-05**	0.168	-0.83	**0.008**	1.80	·
47	cg24898753	11:61735917	FTH1	FTH1	**3.17E-05**	-2.79	**8.47E-05**	0.979	-0.02	**6.51E-04**	2.81	✓
48	cg08240832	1:241791729	OPN3	CHML	**3.21E-05**	7.50	**8.12E-05**	0.129	-1.81	**7.35E-06**	-10.09	✓
49	cg17998333	10:121356721	TIAL1	TIAL1	**3.30E-05**	-1.64	**7.53E-05**	0.441	-0.15	**6.05E-04**	1.41	·
50	cg07170170	12:125672141	-	AACS	**3.37E-05**	-1.63	**7.21E-05**	0.986	0.00	**4.10E-04**	1.59	·
51	cg15474615	3:134370182	KY	KY	**3.47E-05**	-2.93	**7.09E-05**	**0.049**	-1.43	0.055	1.39	·
52	cg01492649	11:86749237	TMEM135	TMEM135	**3.68E-05**	-1.66	**7.65E-05**	0.965	0.01	**4.63E-05**	1.63	·
53	cg09409821	7:98923250	ARPC1A	ARPC1A	**3.70E-05**	-1.62	**8.07E-05**	0.839	0.05	**4.47E-05**	1.72	·
54	cg04108502	12:56360908	CDK2;SILV	CDK2	**3.71E-05**	-2.14	**9.64E-05**	0.200	-0.59	**0.007**	1.54	·
55	cg27175943	16:85334190	-	KIAA0513	**3.72E-05**	-13.55	**9.95E-05**	0.108	-4.44	**0.034**	8.54	✓
56	cg25267526	19:12777743	MAN2B1;MORG1	WDR83	**3.74E-05**	-1.63	**9.77E-05**	0.943	-0.02	**5.29E-05**	1.56	·
57	cg26609202	1:10532686	DFFA	DFFA	**3.74E-05**	-1.49	**7.93E-05**	0.423	-0.11	**2.12E-04**	1.39	·
58	cg07622815	8:22526588	BIN3	BIN3	**4.00E-05**	-1.40	**8.16E-05**	0.579	0.12	**5.37E-05**	1.46	·
59	cg23044270	3:43659607	ANO10	ANO10	**4.12E-05**	6.96	**1.96E-05**	0.410	0.98	**9.30E-04**	-6.72	✓
60	cg14572728	9:98279112	PTCH1	PTCH1	**4.16E-05**	-1.62	**1.02E-04**	0.561	-0.13	**3.92E-04**	1.46	·
61	cg24701270	3:42947990	ZNF662	ZNF662	**4.19E-05**	-2.31	**1.08E-04**	0.400	-0.28	**1.38E-04**	1.96	·
62	cg21196581	3:169780131	GPR160	GPR160	**4.19E-05**	7.09	**5.80E-05**	0.694	0.63	**0.010**	-6.15	·
63	cg07192821	15:45003560	B2M	B2M	**4.21E-05**	-1.56	**7.84E-05**	0.275	-0.21	**1.14E-04**	1.34	·
64	cg12207532	11:8040463	-	TUB	**4.28E-05**	-1.57	**1.14E-04**	0.995	0.00	**1.32E-04**	1.55	·
65	cg24790706	7:157335647	PTPRN2	DNAJB6	**4.30E-05**	7.94	**1.13E-04**	0.240	1.52	**0.008**	-5.90	✓
66	cg09903872	7:157129451	DNAJB6	DNAJB6	**4.45E-05**	-1.48	**1.18E-04**	0.457	-0.16	**5.38E-04**	1.32	·

Table 3.4. cont.

Probe information					5mC (OxBS)			Total modifications (BS)		5hmC (BS - OxBS)		
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	*P*	Δ	*P*	Δ	5hmC Positive probe
67	cg07264124	19:11998686	ZNF69	ZNF69	**4.45E-05**	-1.65	**6.91E-05**	0.293	-0.26	**0.001**	1.39	·
68	cg23436918	7:16685595	BZW2;ANKMY2	ANKMY2	**4.57E-05**	-1.46	**9.85E-05**	0.541	0.10	**1.74E-05**	1.56	·
69	cg18004316	2:208634121	FZD5	FZD5	**4.63E-05**	-1.22	**1.05E-04**	0.646	0.07	**1.61E-04**	1.28	·
70	cg12776966	19:30021276	VSTM2B	VSTM2B	**4.68E-05**	-2.05	**6.77E-05**	0.695	0.17	**0.002**	2.22	·
71	cg06673826	6:111580383	KIAA1919	KIAA1919	**4.76E-05**	-1.13	**1.08E-04**	0.514	0.13	**1.32E-04**	1.26	·
72	cg03958979	6:108486387	NR2E1	NR2E1	**4.77E-05**	-3.07	**1.13E-04**	0.144	-0.75	**0.009**	2.24	·
73	cg09587503	5:134094082	DDX46	DDX46	**4.79E-05**	-1.29	**7.08E-05**	**0.025**	0.41	**2.62E-05**	1.64	·
74	cg25463831	14:58766031	FLJ31306;ARID4A	ARID4A	**4.81E-05**	-1.05	**1.22E-04**	0.200	-0.16	**9.24E-04**	0.90	·
75	cg27460531	17:78791723	RPTOR	CHMP6	**4.92E-05**	-5.48	**9.78E-05**	0.811	0.20	**3.15E-04**	5.53	✓

Table 3.4; Significant neuropathology-associated differentially methylated positions (DMPs) in the EC. *Shown for each probe are chromosomal location (hg19), UCSC annotation, GREAT annotation, with corrected effect size (difference (Δ) in 5mC levels) and corresponding linear regression derived P value between Braak score 0 and Braak score VI after adjusting for the covariates of age and sex, and an additional corresponding P value also correcting for NeuN+ neuronal proportions (P CETS corrected). Shown for all probes are Δ and corresponding P value for total DNA modifications and 5hmC levels alone, with annotation as to whether the mean 5hmC level was >0. All P values <0.05 are shown in bold.*

Probe Information					5hmC (BS - OxBS)				Total modifications (BS)		5mC (OxBS)	
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	5hmC Positive probe	*P*	Δ	*P*	Δ
1	cg15147060	3:88108213	CGGBP1	CGGBP1	**3.94E-07**	1.73	**1.07E-06**	·	0.440	-0.13	**3.65E-09**	-1.92
2	cg25329939	13:98085947	RAP2A	RAP2A	**2.05E-06**	1.56	**4.93E-06**	·	0.074	0.31	**7.73E-06**	-1.29
3	cg17026303	3:38495161	LOC100128640;ACVR2B	ACVR2B	**2.74E-06**	2.39	**7.50E-06**	·	0.948	-0.02	**6.70E-07**	-2.46
4	cg00106685	3:52720133	GNL3;PBRM1	GNL3	**4.04E-06**	1.39	**6.63E-06**	·	0.413	-0.13	**7.80E-08**	-1.45
5	cg01710791	19:19144494	SFRS14;ARMC6	SUGP2	**5.07E-06**	1.41	**1.23E-05**	·	**0.024**	0.38	**3.03E-04**	-1.05
6	cg03160445	3:100120349	LNP1;TOMM70A	TOMM70A	**5.35E-06**	1.50	**1.41E-05**	·	0.680	0.07	**2.46E-05**	-1.48
7	cg05052335	5:130970657	RAPGEF6	RAPGEF6	**6.00E-06**	1.39	**1.48E-05**	·	0.215	-0.17	**9.73E-07**	-1.56
8	cg18070033	2:220025364	NHEJ1	NHEJ1	**6.24E-06**	1.84	**1.38E-05**	·	0.919	0.02	**3.64E-06**	-1.84
9	cg14252850	1:113010353	WNT2B	WNT2B	**6.26E-06**	-9.36	**1.30E-05**	✓	**0.004**	-4.68	**0.009**	4.82
10	cg08240832	1:241791729	OPN3	CHML	**7.35E-06**	-10.09	**1.99E-05**	✓	0.129	-1.81	**3.21E-05**	7.50
11	cg07536920	9:77112326	RORB	RORB	**7.66E-06**	1.91	**2.04E-05**	·	0.655	0.09	**5.99E-05**	-1.67
12	cg11530914	16:67281528	FHOD1;SLC9A5	FHOD1	**7.74E-06**	1.95	**2.09E-05**	·	0.642	0.12	**2.86E-05**	-1.87
13	cg12831866	20:24452036	TMEM90B	SYNDIG1	**8.16E-06**	2.44	**1.54E-05**	·	0.060	0.69	**0.004**	-1.66
14	cg13876163	1:45140495	C1orf228;TMEM53	TMEM53	**8.39E-06**	2.55	**1.45E-05**	·	0.145	0.33	**2.73E-05**	-2.25
15	cg03503087	10:118032626	GFRA1	GFRA1	**1.02E-05**	1.97	**1.79E-05**	·	0.731	0.06	**2.38E-06**	-1.89
16	cg07919466	19:50833956	KCNC3	KCNC3	**1.03E-05**	1.69	**1.68E-05**	·	0.730	-0.09	**8.41E-07**	-1.83
17	cg15481603	15:51977697	SCG3	SCG3	**1.26E-05**	-9.91	**2.98E-05**	✓	0.075	-2.75	**0.003**	6.16
18	cg03263685	2:68480160	PPP3R1	PPP3R1	**1.28E-05**	1.21	**3.25E-05**	·	0.932	0.01	**1.17E-05**	-1.23
19	cg12754733	19:11485436	C19orf39	SWSAP1	**1.29E-05**	1.78	**3.41E-05**	·	0.409	0.16	**5.61E-05**	-1.60
20	cg08888916	11:118016368	SCN4B	SCN4B	**1.50E-05**	1.72	**3.62E-05**	·	0.841	0.04	**5.44E-05**	-1.62
21	cg12144476	21:27541894	APP	APP	**1.61E-05**	1.44	**4.12E-05**	·	0.659	0.08	**4.85E-06**	-1.39
22	cg19370451	1:152386561	CRNN	CRNN	**1.63E-05**	-5.78	**3.42E-05**	✓	**0.002**	-3.04	**0.029**	2.73
23	cg23436918	7:16685595	BZW2;ANKMY2	ANKMY2	**1.74E-05**	1.56	**4.52E-05**	·	0.541	0.10	**4.57E-05**	-1.46
24	cg06174962	10:44351762	-	ZNF32	**1.88E-05**	-8.83	**4.81E-05**	✓	**0.024**	-2.39	**2.72E-04**	6.43
25	cg19005236	8:1571105	DLGAP2	DLGAP2	**1.93E-05**	-10.74	**2.48E-05**	·	0.341	-1.82	**2.68E-04**	8.76
26	cg16498314	16:30366841	CD2BP2	CD2BP2	**1.99E-05**	1.36	**3.81E-05**	·	0.271	0.18	**2.37E-04**	-1.11
27	cg05341384	16:86331661	-	FOXF1	**2.01E-05**	4.41	**4.84E-05**	·	0.875	-0.10	**4.15E-06**	-4.92
28	cg23665824	2:80530701	CTNNA2;LRRTM1	LRRTM1	**2.10E-05**	1.11	**4.98E-05**	✓	0.054	0.39	**6.15E-04**	-0.68
29	cg04149978	1:10764896	CASZ1	CASZ1	**2.17E-05**	-5.77	**2.28E-05**	·	0.310	-1.27	**0.003**	4.46
30	cg08318505	11:133903913	LOC100128239	JAM3	**2.26E-05**	2.20	**4.00E-05**	·	**0.001**	0.83	**0.003**	-1.31
31	cg10663655	7:66309627	LOC729156	TMEM248	**2.27E-05**	1.77	**5.56E-05**	·	0.061	0.28	**2.45E-04**	-1.49

Table 3.5.

Probe Information					5hmC (BS - OxBS)				Total modifications (BS)		5mC (OxBS)	
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	5hmC Positive probe	*P*	Δ	*P*	Δ
32	cg02493644	7:84816314	-	SEMA3D	**2.43E-05**	1.60	**5.47E-05**	·	0.073	0.36	**3.34E-04**	-1.22
33	cg09587503	5:134094082	DDX46	DDX46	**2.62E-05**	1.64	**5.31E-05**	·	**0.025**	0.41	**4.79E-05**	-1.29
34	cg08182454	19:46529189	-	PGLYRP1	**2.67E-05**	-10.76	**6.58E-05**	✓	0.106	-2.73	**0.001**	7.40
35	cg20066612	1:160050948	KCNJ9	KCNJ9	**2.70E-05**	5.84	**6.47E-05**	✓	**6.79E-06**	3.25	**0.027**	-2.89
36	cg13009654	5:137802252	EGR1	EGR1	**2.71E-05**	3.94	**3.00E-05**	·	0.070	1.01	**8.30E-04**	-2.91
37	cg16933762	3:77147109	ROBO2	-	**2.80E-05**	-10.61	**7.11E-05**	·	0.493	-1.12	**2.77E-05**	10.36
38	cg25954627	10:8090846	-	GATA3	**3.31E-05**	3.38	**7.17E-05**	·	0.484	0.29	**3.39E-04**	-3.00
39	cg18494448	17:16492891	-	ZNF287	**3.41E-05**	1.59	**9.07E-05**	·	0.150	0.26	**1.75E-04**	-1.40
40	cg05503627	12:26267490	-	BHLHE41	**3.47E-05**	0.88	**4.10E-05**	✓	**0.002**	0.54	0.197	-0.23
41	cg20894246	22:41682263	RANGAP1	RANGAP1	**3.48E-05**	2.24	**9.19E-05**	·	0.108	0.38	**4.97E-04**	-1.86
42	cg03436461	17:13504687	HS3ST3A1	HS3ST3A1	**3.68E-05**	2.08	**9.77E-05**	·	0.565	0.17	**1.39E-04**	-1.91
43	cg17337106	8:80577176	STMN2	STMN2	**3.80E-05**	-6.70	**9.62E-05**	·	0.075	-2.19	**8.36E-04**	3.93
44	cg17848546	1:209605304	LOC642587;MIR205	CAMK1G	**3.82E-05**	-7.82	**9.14E-05**	✓	**0.035**	-2.68	**0.007**	4.61
45	cg09504196	7:148823235	ZNF425;ZNF398	ZNF425	**3.85E-05**	1.58	**7.87E-05**	·	0.113	0.29	**1.54E-04**	-1.26
46	cg15624314	1:226251003	H3F3A;LOC440926	H3F3A	**4.30E-05**	1.59	**9.50E-05**	·	0.320	-0.25	**6.14E-07**	-1.84
47	cg09409821	7:98923250	ARPC1A	ARPC1A	**4.47E-05**	1.72	**9.14E-05**	·	0.839	0.05	**3.70E-05**	-1.62
48	cg19878762	22:23484297	RTDR1	RTDR1	**4.49E-05**	1.49	**1.08E-04**	·	0.126	0.28	**4.17E-04**	-1.24
49	cg22026192	12:124944377	NCOR2	NCOR2	**4.51E-05**	-8.41	**1.14E-04**	✓	**0.013**	-2.91	**0.003**	5.56
50	cg01492649	11:86749237	TMEM135	TMEM135	**4.63E-05**	1.63	**9.62E-05**	·	0.965	0.01	**3.68E-05**	-1.66
51	cg20966551	19:12949060	MAST1	MAST1	**4.66E-05**	1.59	**1.02E-04**	·	0.231	0.23	**9.07E-05**	-1.39
52	cg14546505	9:102863959	INVS	INVS	**4.69E-05**	-7.18	**6.03E-05**	✓	0.089	-2.41	**0.020**	4.28
53	cg17158913	1:10764886	CASZ1	CASZ1	**4.78E-05**	-8.74	**9.33E-05**	·	0.366	-1.16	**8.94E-04**	7.56
54	cg14717061	11:126138971	FOXRED1;SRPR	FOXRED1	**4.89E-05**	1.28	**7.78E-05**	·	0.562	0.08	**7.13E-05**	-1.20
55	cg07379167	10:131769074	-	EBF3	**4.89E-05**	2.45	**1.31E-04**	·	0.062	-0.77	**5.45E-05**	-3.02
56	cg07269003	3:142681516	PAQR9	PAQR9	**4.90E-05**	-8.72	**6.60E-05**	·	**0.021**	-3.64	**0.009**	4.74
57	cg24432048	1:3827651	LOC100133612	C1orf174	**4.97E-05**	1.91	**6.87E-05**	·	**0.050**	0.55	**0.001**	-1.29

Table 3.5; Significant neuropathology-associated differentially hydroxymethylated positions (DHPs) in the EC. *Shown for each probe are chromosomal location (hg19), UCSC annotation, GREAT annotation, with corrected effect size (difference (Δ) in 5hmC levels) and corresponding linear regression derived P value between Braak score 0 and Braak score VI after adjusting for the covariates of age and sex, and an additional corresponding P value correcting for NeuN+ neuronal proportions (P CETS corrected), and annotation as to whether the mean 5hmC level was >0. Shown for all probes are Δ and corresponding P value for total DNA modifications and 5mC levels alone. All P values <0.05 are shown in bold.*

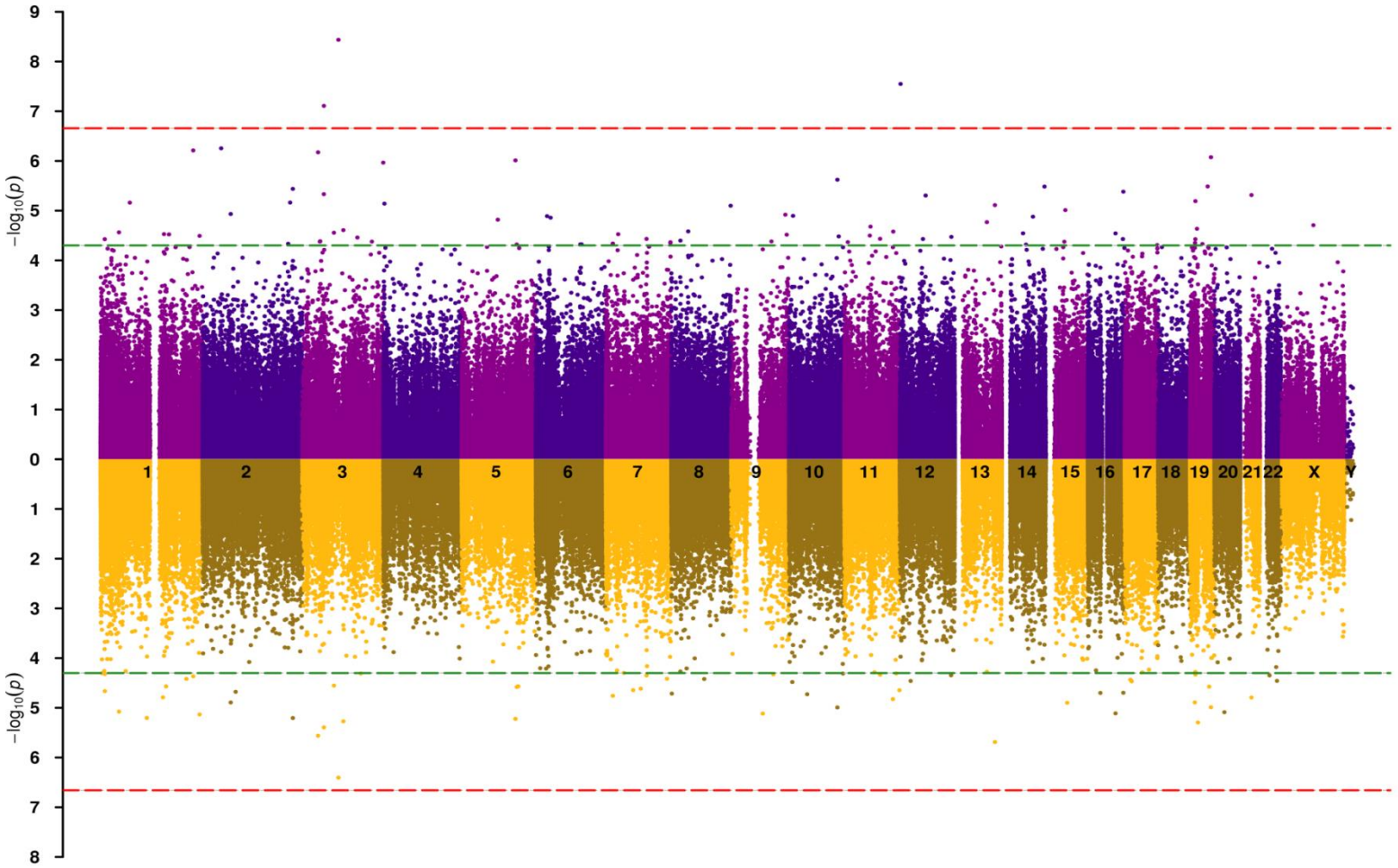


Figure 3.7; Alterations in total DNA modifications, DNA methylation and DNA hydroxymethylation are associated with AD neuropathology in the EC. *(**A**) A Manhattan plot of association between DNA methylation (purple, above X axis) and DNA hydroxymethylation (yellow, below X axis) with Braak stage in the EC highlights associations at loci across the genome. Using linear regression we identified three DMPs reaching experiment wide significance (red line; $P < 2.2 \times 10^{-7}$), and a further 72 DMPs reaching our nominal significance threshold (green line; $P < 5 \times 10^{-5}$). Although no DHPs reached experiment wide significance, we identified 57 DHPs at our nominal significance threshold.*

Probe information					5mC (OxBS)			Total modifications (BS)		5hmC (BS - OxBS)		
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	*P*	Δ	*P*	Δ	5hmC Positive probe
1	cg04868540	11:134263237	B3GAT1	B3GAT1	**7.05E-07**	7.69	8.14E-07	8.31E-01	0.23	3.99E-05	-7.42	✓
2	cg23448978	7:51209365	COBL	COBL	**5.10E-06**	-1.67	4.45E-06	2.32E-02	-0.58	1.11E-02	1.09	✓
3	cg12898370	12:24737779	C12orf67	BCAT1	**1.13E-05**	5.98	4.93E-06	6.04E-01	-0.91	1.85E-03	-6.78	✓
4	cg21548096	12:69326897	CPM	CPM	**2.21E-05**	0.80	2.48E-05	5.37E-01	0.11	5.96E-03	-0.68	✓
5	cg08156809	7:74071310	GTF2I	GTF2I	**3.23E-05**	5.78	3.10E-05	4.89E-01	-0.55	3.32E-05	-6.39	▪
6	cg21858113	11:118015340	SCN4B	SCN4B	**4.09E-05**	5.05	2.14E-05	2.20E-01	1.22	1.84E-04	-3.88	✓
7	cg02592271	2:27665507	KRTCAP3	KRTCAP3	**4.12E-05**	-9.06	1.96E-05	6.36E-02	-3.12	1.54E-02	5.87	✓

Table 3.6; Significant neuropathology-associated differentially methylated positions (DMPs) in the CER. *Shown for each probe are chromosomal location (hg19), UCSC annotation, GREAT annotation, with corrected effect size (difference (Δ) in 5mC levels) and corresponding linear regression derived P value between Braak score 0 and Braak score VI after adjusting for the covariates of age and sex, and an additional corresponding P value also correcting for NeuN+ neuronal proportions (P CETS corrected). Shown for all probes are Δ and corresponding P value for total DNA modifications and 5hmC levels alone, with annotation as to whether the mean 5hmC level was >0. All P values <0.05 are shown in bold.*

Probe Information					5hmC (BS - OxBS)				Total modifications (BS)		5mC (OxBS)	
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	5hmC Positive probe	*P*	Δ	*P*	Δ
1	cg09214953	12:111800702	FAM109A	FAM109A	**4.40E-06**	-5.70	6.26E-06	✓	3.31E-03	-2.40	4.51E-03	3.33
2	cg01385157	15:89421347	HAPLN3	HAPLN3	**1.35E-05**	-8.82	1.89E-05	✓	1.72E-01	-5.63	3.39E-01	3.40
3	cg04116354	1:26003643	MAN1C1	MAN1C1	**1.41E-05**	10.67	1.72E-05	✓	1.12E-02	3.96	1.55E-03	-6.71
4	cg10231603	1:155829206	SYT11	GON4L	**2.11E-05**	-2.75	1.39E-05	▪	1.78E-01	-0.59	1.36E-03	2.17
5	cg14289985	19:57018848	ZNF471	ZNF471	**2.61E-05**	1.53	3.16E-05	✓	4.17E-05	1.07	1.61E-01	-0.44
6	cg17606548	5:66123702	MAST4	MAST4	**2.66E-05**	-8.44	2.89E-05	✓	1.76E-01	-1.87	2.20E-04	6.70
7	cg16061228	11:118014547	SCN4B	SCN4B	**2.89E-05**	-5.40	4.03E-05	✓	2.50E-01	-1.22	1.29E-02	4.21
8	cg08156809	7:74071310	GTF2I	GTF2I	**3.32E-05**	-6.39	2.43E-05	▪	4.89E-01	-0.55	3.23E-05	5.78
9	cg23180780	2:12260486	-	LPIN1	**3.39E-05**	-5.74	4.16E-05	▪	5.16E-02	-1.94	9.08E-03	3.83
10	cg00732538	16:87635674	JPH3	JPH3	**3.41E-05**	-2.78	3.25E-05	▪	5.08E-01	-0.19	6.53E-05	2.57
11	cg10461218	12:124144432	GTF2H3	TCTN2	**3.85E-05**	-5.54	3.81E-05	✓	1.59E-01	-1.20	9.88E-05	4.33
12	cg04868540	11:134263237	B3GAT1	B3GAT1	**3.99E-05**	-7.42	5.22E-05	✓	8.31E-01	0.23	7.05E-07	7.69

Table 3.7; Significant neuropathology-associated differentially hydroxymethylated positions (DHPs) in the CER. *Shown for each probe are chromosomal location (hg19), UCSC annotation, GREAT annotation, with corrected effect size (difference (Δ) in 5hmC levels) and corresponding linear regression derived P value between Braak score 0 and Braak score VI after adjusting for the covariates of age and sex, and an additional corresponding P value correcting for neuron/glia proportions (P CETS corrected), and annotation as to whether the mean 5hmC level was >0. Shown for all probes are Δ and corresponding P value for total DNA modifications and 5mC levels alone. All P values <0.05 are shown in bold.*

3.4.4 Neuropathology-associated DMPs and DHPs in the EC feature in key biological pathways

We performed a pathway analysis of the genes annotated to DMPs and DHPs in the EC that reached our nominal significance threshold ($P < 5 \times 10^{-5}$). This *P* value threshold was used to limit the number of associated genes included, in an attempt to prevent the identification of false positive pathways. (Table 3.8 and Table 3.9, respectively). As this analysis represented exploratory work all significant pathways ($P < 0.005$) are presented. Relevant pathways of interest included axon midline choice point recognition" and "axon choice point recognition", both of which were enriched for DMPs and DHPs (GO:0016199: 5mC $P = 2.30 \times 10^{-4}$, 5hmC $P = 1.30 \times 10^{-4}$; GO:0016198: 5mC $P = 4.20 \times 10^{-4}$, 5hmC $P = 2.55 \times 10^{-4}$), "regulation of synapse organisation" (GO:0050807: 5hmC $P = 9.11 \times 10^{-4}$), "synapse organisation" (GO:0050808: 5hmC $P = 0.001$) and "synapse assembly" (GO:0007416: 5hmC $P = 0.002$), which were all enriched for DHPs.

GO Term	GO Name	Type	No. Genes in Pathway	No. Test List in Pathway	*P* Genes in Test List
GO:0060715	syncytiotrophoblast cell differentiation involved in labyrinthine layer development	BP	3	2	8.26E-05
GO:0016199	axon midline choice point recognition	BP	4	2	2.30E-04
GO:0060717	chorion development	BP	7	2	3.92E-04
GO:0000805	X chromosome	CC	7	2	4.15E-04
GO:0016198	axon choice point recognition	BP	6	2	4.20E-04
GO:0005773	vacuole	CC	1157	13	5.44E-04
GO:1903867	extraembryonic membrane development	BP	9	2	6.07E-04
GO:0021772	olfactory bulb development	BP	32	3	6.14E-04
GO:0021988	olfactory lobe development	BP	33	3	6.86E-04
GO:0042246	tissue regeneration	BP	50	3	8.39E-04
GO:0002726	positive regulation of T cell cytokine production	BP	14	2	9.35E-04
GO:0044440	endosomal part	CC	416	7	0.001
GO:0044437	vacuolar part	CC	684	9	0.001
GO:0002724	regulation of T cell cytokine production	BP	19	2	0.001
GO:0045893	positive regulation of transcription, DNA-templated	BP	1301	14	0.001
GO:1903508	positive regulation of nucleic acid-templated transcription	BP	1301	14	0.001
GO:1902680	positive regulation of RNA biosynthetic process	BP	1322	14	0.002
GO:0021889	olfactory bulb interneuron differentiation	BP	12	2	0.002
GO:0051254	positive regulation of RNA metabolic process	BP	1367	14	0.002
GO:0038023	signaling receptor activity	MF	1227	10	0.002
GO:0030033	microvillus assembly	BP	16	2	0.002
GO:0009880	embryonic pattern specification	BP	58	3	0.003
GO:0031904	endosome lumen	CC	22	2	0.003
GO:0030139	endocytic vesicle	CC	246	5	0.003
GO:0002369	T cell cytokine production	BP	25	2	0.003
GO:0034219	carbohydrate transmembrane transport	BP	26	2	0.003
GO:0035994	response to muscle stretch	BP	19	2	0.003
GO:0043235	receptor complex	CC	318	6	0.003
GO:0004888	transmembrane signaling receptor activity	MF	1119	9	0.003
GO:0030666	endocytic vesicle membrane	CC	147	4	0.003
GO:0060644	mammary gland epithelial cell differentiation	BP	16	2	0.003
GO:0033267	axon part	CC	206	5	0.003
GO:0030673	axolemma	CC	14	2	0.003

Table 3.8.

GO Term	GO Name	Type	No. Genes in Pathway	No. Test List in Pathway	*P* Genes in Test List
GO:0006357	regulation of transcription from RNA polymerase II promoter	BP	1761	16	0.003
GO:0006366	transcription from RNA polymerase II promoter	BP	1947	17	0.003
GO:0010628	positive regulation of gene expression	BP	1616	15	0.003
GO:0008093	cytoskeletal adaptor activity	MF	15	2	0.003
GO:0044304	main axon	CC	56	3	0.003
GO:0005768	endosome	CC	771	9	0.004
GO:0019731	antibacterial humoral response	BP	40	2	0.004
GO:0060713	labyrinthine layer morphogenesis	BP	21	2	0.004
GO:0021872	forebrain generation of neurones	BP	63	3	0.004
GO:0007176	regulation of epidermal growth factor-activated receptor activity	BP	25	2	0.004
GO:0098659	inorganic cation import into cell	BP	25	2	0.004
GO:0099587	inorganic ion import into cell	BP	25	2	0.004
GO:0010008	endosome membrane	CC	386	6	0.004
GO:0002711	positive regulation of T cell mediated immunity	BP	31	2	0.004
GO:0002720	positive regulation of cytokine production involved in immune response	BP	31	2	0.005
GO:0099600	transmembrane receptor activity	MF	1173	9	0.005
GO:0030332	cyclin binding	MF	20	2	0.005
GO:0019730	antimicrobial humoral response	BP	45	2	0.005
GO:0000803	sex chromosome	CC	27	2	0.005
GO:0032528	microvillus organization	BP	24	2	0.005

Table 3.8; Pathways enriched with neuropathology-associated DMPs in EC. *Abbreviations: Gene Ontology (GO), Biological Process (BP), Cellular Component (CC), Molecular Function (MF). Shown are pathways were P < 0.005. Data is organized with the most significant pathways at the top of the list.*

GO Term	GO Name	Type	No. Genes in Pathway	No. Test List in Pathway	*P* Genes in Test List
GO:0050803	regulation of synapse structure or activity	BP	115	5	8.24E-05
GO:0016199	axon midline choice point recognition	BP	4	2	1.30E-04
GO:0016198	axon choice point recognition	BP	6	2	2.55E-04
GO:0050807	regulation of synapse organization	BP	111	4	9.11E-04
GO:0050808	synapse organization	BP	229	5	0.001
GO:0030033	microvillus assembly	BP	16	2	0.002
GO:0007416	synapse assembly	BP	135	4	0.002
GO:0033233	regulation of protein sumoylation	BP	20	2	0.002
GO:0030424	axon	CC	398	6	0.003
GO:0032528	microvillus organization	BP	24	2	0.003
GO:0051963	regulation of synapse assembly	BP	78	3	0.004

Table 3.9; Pathways enriched with neuropathology-associated DHPs in EC*. Abbreviations: Gene Ontology (GO), Biological Process (BP), Cellular Component (CC), Molecular Function (MF). Shown are pathways were P < 0.005. Data is organized with the most significant pathways at the top of the list. Data is organized with the most significant pathways at the top of the list.*

3.4.5 Neuropathology-associated changes in 5mC and 5hmC are seen across gene features

Next, we compared the genetic location of the most significant hyper- or hypo-(hydroxy)methylated neuropathology-associated loci in the EC ($P < 5 \times 10^{-5}$), to all nominally-significant neuropathology-associated DMPs or DHPs ($P < 0.05$), with relation to gene features that have been described by Slieker et al. (2013) (Figures 3.8-3.9; Tables 3.10-3.11). We demonstrated an enrichment for the most significant DMPs and DHPs across all gene features, most strikingly Braak-associated hypomethylation and hyperhydroxymethylation appears to be enriched in the CGIs of proximal promotors ($P < 1 \times 10^{-330}$).

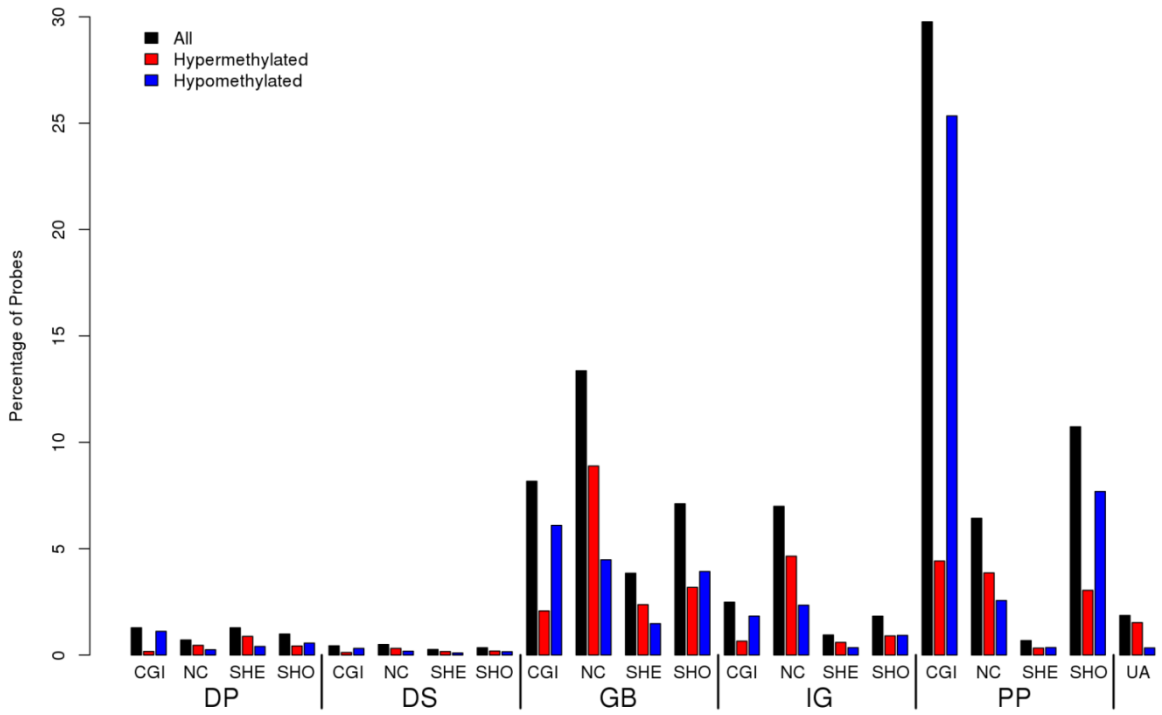


Figure 3.8: The most significant neuropathology-associated DMPs are enriched at specific genomic features in the EC*. Taking the genomic location of neuropathology-associated DMPs above our significance threshold (P < 0.05), we used a Fisher's exact test to examine whether these hypermethylated or hypomethylated loci are enriched within specific genomic features when compared to all Braak-associated DMPs (P < 0.05) with results of the enrichments tests shown in Supplementary Table 9. Abbreviations: CpG island (CGI), outside CpG island (nonCGI or NC), CpG island shelf (SHE), CpG island shore (SHO), distal promoter (DP), downstream (DS), gene body (GB), intergenic (IG), proximal promoter (PP), unannotated (UA).*

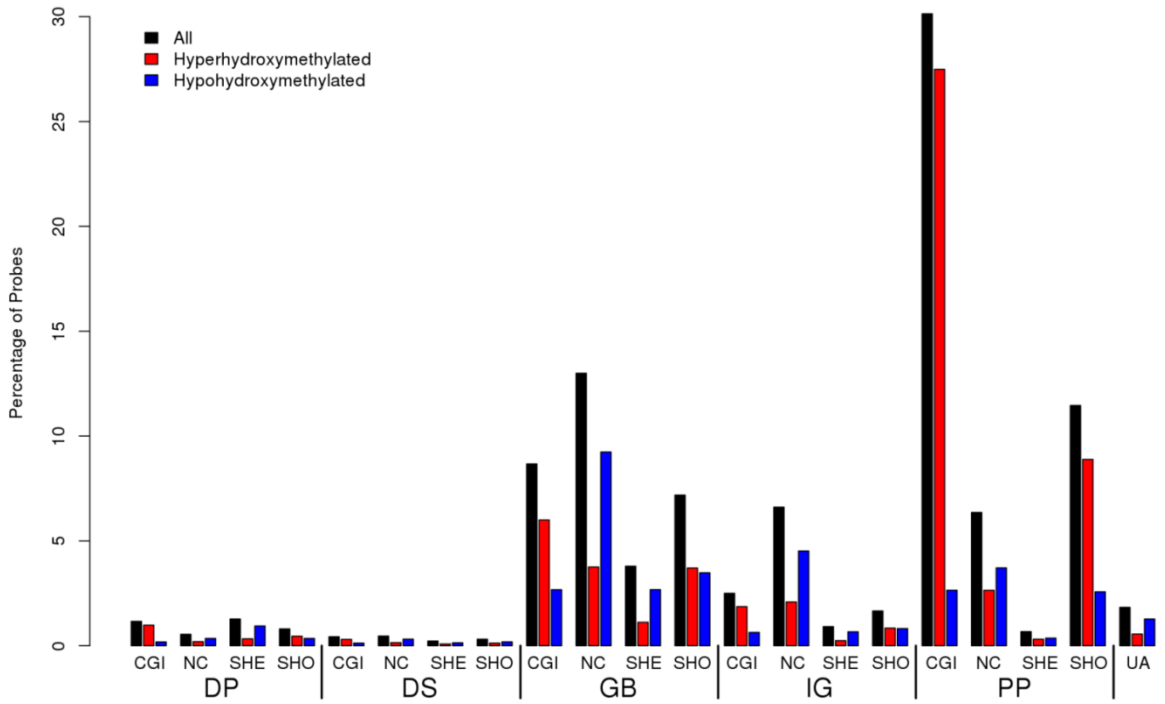


Figure 3.9; The most significant neuropathology-associated DHPs are enriched at specific genomic features in the EC*. Taking the genomic location of neuropathology-associated DHPs above our significance threshold (P < 0.05), we used a Fisher's exact test to examine whether these hyperhydroxymethylated or hypohydroxymethylated loci are enriched within specific genomic features when compared to all Braak-associated DHPs (P <0.05), with results of the enrichments tests shown in Supplementary Table 10. Abbreviations: CpG island (CGI), outside CpG island (nonCGI or NC), CpG island shelf (SHE), CpG island shore (SHO), distal promoter (DP), downstream (DS), gene body (GB), intergenic (IG), proximal promoter (PP), unannotated (UA).*

		Significant Hypermethylation	Direction	*P* value	Significant Hypomethylation	Direction	*P* value
DP	**CGI**	✓	-	7.01E-29	✓	+	2.08E-42
	NC	·			✓	-	1.14E-07
	SHE	·			✓	-	1.75E-21
	SHO	✓	-	7.61E-04	·		
DS	**CGI**	✓	-	1.50E-07	✓	+	0.008
	NC	·			✓	-	1.72E-06
	SHE	·			✓	-	0.003
	SHO	·			✓	-	0.001
GB	**CGI**	✓	-	3.65E-108	✓	+	1.56E-64
	NC	✓	+	1.08E-05	✓	-	2.55E-168
	SHE	·			✓	-	6.84E-41
	SHO	✓	-	1.04E-30	✓	+	3.43E-05
IG	**CGI**	✓	-	1.04E-28	✓	+	5.60E-22
	NC	✓	+	1.46E-04	✓	-	5.35E-84
	SHE	·			✓	-	3.92E-09
	SHO	✓	-	1.74E-04	✓	+	0.001
PP	**CGI**	✓	-	4.94E-324	✓	+	<1.00E-330
	NC	✓	+	0.012	✓	-	1.93E-29
	SHE	✓	-	0.004	✓	+	0.037
	SHO	✓	-	6.10E-128	✓	+	1.48E-49
	UA	✓	+	8.35E-08	✓	-	3.47E-73

Table 3.10; Table highlighting specific genomic features enriched for neuropathology-associated DMPs. *Taking the genomic location of neuropathology-associated DMPs above our significance threshold ($P < 5 \times 10^{-5}$), we used a Fisher's exact test to examine whether hypermethylated or hypomethylated loci are enriched within specific genomic features compared to all Braak-associated probes ($P < 0.05$). Shown for each Slieker genomic feature is whether we observed significant disease-associated hypermethylation or hypomethylation, direction of change and corresponding P value. Abbreviations: CpG island (CGI), outside CpG island (nonCGI or NC), CpG island shelf (SHE), CpG island shore (SHO), distal promoter (DP), downstream (DS), gene body (GB), intergenic (IG), proximal promoter (PP), unannotated (UA).*

		Significant Hyperhydroxymethylation	Direction	*P* value	Significant Hypohydroxymethylation	Direction	*P* value
DP	**CGI**	✓	+	1.62E-10	✓	-	3.30E-33
	NC	✓	-	6.28E-19	✓	+	1.25E-05
	SHE	✓	-	4.54E-41	·		
	SHO	✓	+	2.86E-07	✓	-	1.16E-14
DS	**CGI**	·			✓	-	3.93E-10
	NC	✓	-	2.89E-14	✓	+	0.037
	SHE	✓	-	1.78E-06	✓	+	0.028
	SHO	✓	-	3.09E-09	✓	+	3.21E-04
GB	**CGI**	✓	+	9.97E-09	✓	-	1.15E-100
	NC	✓	-	4.94E-324	✓	+	1.69E-08
	SHE	✓	-	2.64E-111	✓	-	1.57E-05
	SHO	✓	+	7.09E-43	✓	-	3.43E-55
IG	**CGI**	✓	+	1.59E-05	✓	-	1.64E-46
	NC	✓	-	6.27E-165	✓	+	3.19E-09
	SHE	✓	-	3.52E-27	·		
	SHO	✓	+	4.17E-17	✓	-	8.78E-19
PP	**CGI**	✓	+	<1.00E-330	✓	-	4.94E-324
	NC	✓	-	7.26E-61	✓	+	2.10E-11
	SHE	✓	-	3.22E-08	✓	+	4.17E-05
	SHO	✓	+	3.76E-26	✓	-	1.29E-257
	UA	✓	-	2.83E-53	✓	+	7.22E-04

Table 3.11; Table highlighting specific genomic features enriched for neuropathology-associated DHPs. *Taking the genomic location of neuropathology-associated DHPs above our significance threshold ($P < 5 \times 10^{-5}$), we used a Fisher's exact test to examine whether hyperhydroxymethylated or hypohydroxymethylated loci are enriched within specific genomic features compared to all Braak-associated probes ($P < 0.05$). Shown for each Slieker genomic feature is whether we observed significant disease-associated hyperhydroxymethylation or hypohydroxymethylation, direction of change and corresponding P value. Abbreviations: CpG island (CGI), outside CpG island (nonCGI or NC), CpG island shelf (SHE), CpG island shore (SHO), distal promoter (DP), downstream (DS), gene body (GB), intergenic (IG), proximal promoter (PP), unannotated (UA).*

3.4.6 A number of loci display significant negative correlations in neuropathology-associated DNA methylation and hydroxymethylation changes

Next, we sought to explore whether there was any overlap in disease-associated DMoPs, DMPs and DHPs. We compiled a list of loci that showed a disease-associated alteration in more than one analysis ($P < 5 \times 10^{-5}$) (Figure 3.10; Table 3.12). We identified 14 sites (annotated to 14 genes) that showed significant changes across more than one analysis (total modifications, 5mC alone and 5hmC alone) (Figure 3.11). Of the 14 CpGs nominated, 13 showed no nominally significant ($P > 0.05$) difference in total DNA modifications. Although we did observe neuropathology-associated differences in both 5mC and 5hmC, with in each instance the levels of both modifications being highly negatively correlated with each other (Figure 3.11) and a significant interaction between the two modifications (Figure 3.12). Of particular interest are differences at a site annotated to the *APP* gene (cg12144476). In the current study we saw a decrease 5mC in *APP* with Braak stage (Figure 3.11N; Effect size (difference (∆) Braak 0 to Braak VI) = -1.39, $P = 4.85 \times 10^{-06}$), with a concurrent parallel increase in 5hmC (Figure 3.11N; ∆ = 1.44, $P = 1.61 \times 10^{-05}$) demonstrating a significant interaction between the two modifications (Figure 3.12M; $P = 1.77 \times 10^{-10}$) and no net difference in total DNA modifications.

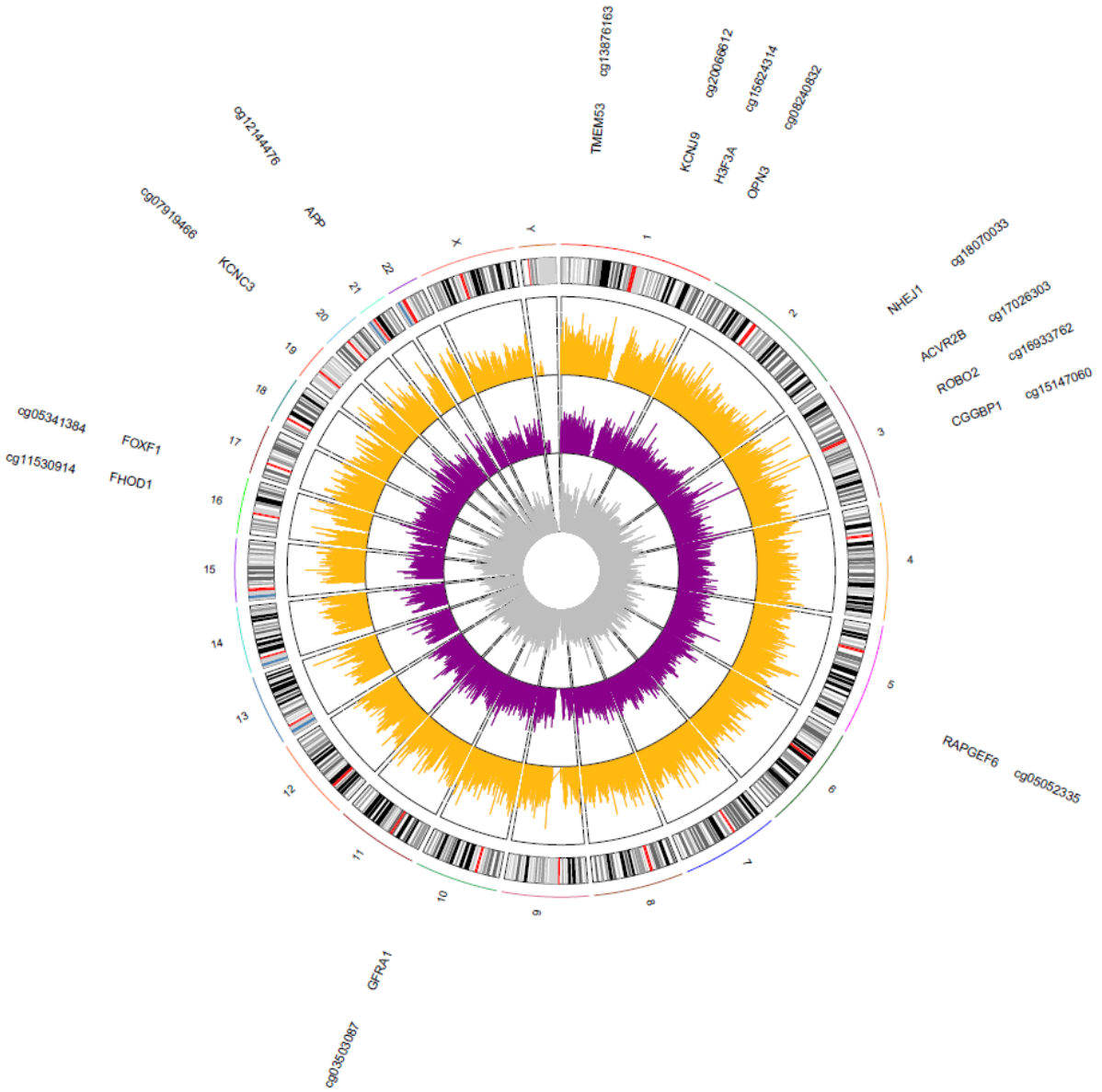


Figure 3.10; Alterations in total DNA modifications, DNA methylation and DNA hydroxymethylation are associated with AD neuropathology in the EC. *A circos-plot highlighting Braak-associated DMoPs (grey–inner circle), DMPs (purple–middle circle) and DHPs (yellow – outer circle). Loci that showed a disease-associated epigenetic alteration in more than one analysis ($P < 5 \times 10^{-5}$) are inscribed around the outside.*

			DMoP (BS)			DMP (OxBS)			DHP (BS - OxBS)			
Gene name	**Probe**	**Position**	***P***	**Δ**	***P* CETS corrected**	***P***	**Δ**	***P* CETS corrected**	***P***	**Δ**	***P* CETS corrected**	***5hmC* Positive probe**
KCNJ9	cg20066612	1:160050948	**$6.79x10^{-06}$**	**3.25**	**$2.73x10^{-05}$**	**0.027**	**-2.89**	**0.033**	**$2.70x10^{-05}$**	**5.84**	**$6.47X10^{-05}$**	✓
H3F3A	cg15624314	1:226251003	0.32	-0.25	0.45	**$6.14x10^{-07}$**	**-1.84**	**$1.68x10^{-06}$**	**$4.30x10^{-05}$**	**1.59**	**$9.50X10^{-05}$**	·
OPN3	cg08240832	1:241791729	0.129	-1.81	0.097	**$3.21x10^{-05}$**	**7.5**	**$8.12x10^{-05}$**	**$7.35x10^{-06}$**	**-10.09**	**$1.99X10^{-05}$**	✓
TMEM53	cg13876163	1:45140495	0.145	0.33	0.178	**$2.73x10^{-05}$**	**-2.25**	**$4.28x10^{-05}$**	**$8.39x10^{-06}$**	**2.55**	**$1.45X10^{-05}$**	·
NHEJ1	cg18070033	2:220025364	0.919	0.02	0.848	**$3.64x10^{-06}$**	**-1.84**	**$8.65x10^{-06}$**	**$6.24x10^{-06}$**	**1.84**	**$1.38X10^{-05}$**	·
ACVR2B	cg17026303	3:38495161	0.948	-0.02	0.976	**$6.70x10^{-07}$**	**-2.46**	**$1.84x10^{-06}$**	**$2.74x10^{-06}$**	**2.39**	**$7.50X10^{-06}$**	·
ROBO2	cg16933762	3:77147109	0.493	-1.12	0.386	**$2.77x10^{-05}$**	**10.36**	**$7.41x10^{-05}$**	**$2.80x10^{-05}$**	**-10.61**	**$7.11X10^{-05}$**	·
CGGBP1	cg15147060	3:88108213	0.44	-0.13	0.536	**$3.65x10^{-09}$**	**-1.92**	**$1.01x10^{-08}$**	**$3.94x10^{-07}$**	**1.73**	**$1.07X10^{-06}$**	·
RAPGEF6	cg05052335	5:130970657	0.215	-0.17	0.217	**$9.73x10^{-07}$**	**-1.56**	**$2.38x10^{-06}$**	**$6.00x10^{-06}$**	**1.39**	**$1.48X10^{-05}$**	·
GFRA1	cg03503087	10:118032626	0.731	0.06	0.765	**$2.38x10^{-06}$**	**-1.89**	**$4.50x10^{-06}$**	**$1.02x10^{-05}$**	**1.97**	**$1.79X10^{-05}$**	·
FHOD1	cg11530914	16:67281528	0.642	0.12	0.68	**$2.86x10^{-05}$**	**-1.87**	**$7.39x10^{-05}$**	**$7.74x10^{-06}$**	**1.95**	**$2.09X10^{-05}$**	·
FOXF1	cg05341384	16:86331661	0.875	-0.1	0.948	**$4.15x10^{-06}$**	**-4.92**	**$1.10x10^{-05}$**	**$2.01x10^{-05}$**	**4.41**	**$4.84X10^{-05}$**	·
KCNC3	cg07919466	19:50833956	0.73	-0.09	0.986	**$8.41x10^{-07}$**	**-1.83**	**$2.28x10^{-06}$**	**$1.03x10^{-05}$**	**1.69**	**$1.68X10^{-05}$**	·
APP	cg12144476	21:27541894	0.659	0.08	0.562	**$4.85x10^{-06}$**	**-1.39**	**$9.79x10^{-06}$**	**$1.61x10^{-05}$**	**1.44**	**$4.12X10^{-05}$**	·

Table 3.12; DMoPs, DMPs and DHPs with a neuropathology-associated difference in levels above our significance threshold ($P < 5.0 \times 10^{-5}$) in more than one comparison. *Shown for each probe are chromosomal location (hg19), Gene annotation (UCSC/GREAT), with corrected effect size (difference (Δ) between Braak score 0 and Braak score VI) and corresponding P value after adjusting for the covariates of age and sex, and an additional P value after correcting for NeuN+ neuronal proportions (P CETS corrected) for total DNA modifications (BS), 5mC (OxBS) and 5hmC (BS – OxBS), with annotation as to whether the mean 5hmC level was >0. All P values <0.05 are shown in bold.*

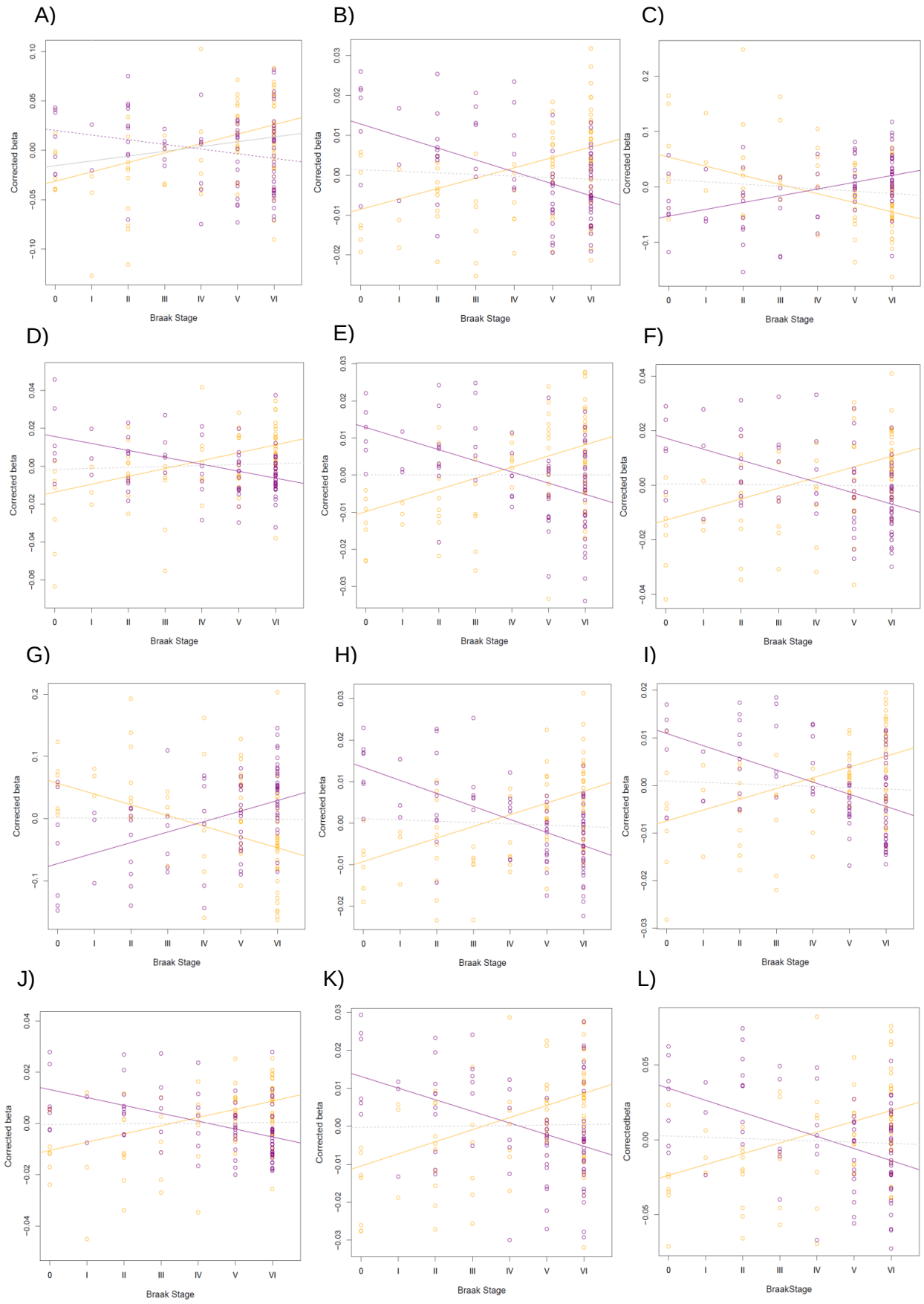

A)
B)
C)
D)
E)
F)
G)
H)
I)
J)
K)
L)
Corrected beta
Corrected beta
Correctedbeta
Braak Stage
BraakStage
0
I
II
III
IV
V
VI

Figure 3.11.

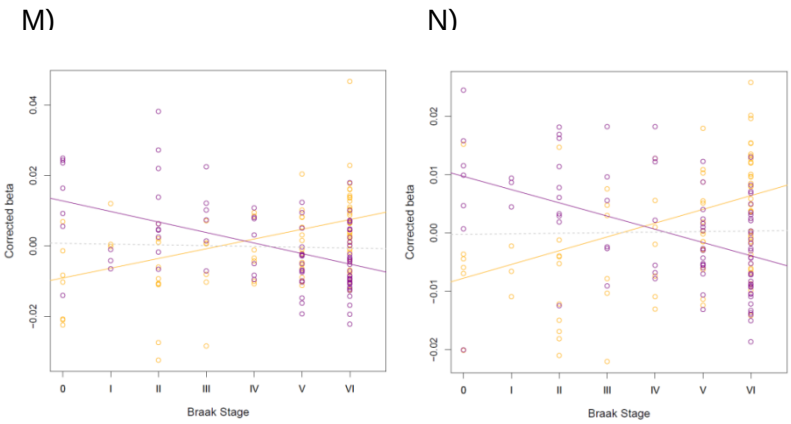


Figure 3.11; Scatter plot of beta values against Braak stage for the 14 probes that feature in Table 3.12. *Shown are beta values, corrected for age and sex, for (A) cg20066612 (KCNJ9), (B) cg15624314 (H3F3A), (C) cg08240832 (OPN3), (D) cg13876163 (TMEM53), (E) cg18070033 (NHEJ1), (F) cg17026303 (ACVR2B), (G) cg16933762 (ROBO2), (H) cg15147060 (CGGBP1), (I) cg05052335 (RAPGEF6), (J) cg03503087 (GFRA1), (K) cg11530914 (FHOD1), (L) cg05341384 (FOXF1), (M) cg07919466 (KCNC3) and (N) cg12144476 (APP). Shown are beta values corrected for age and sex (Y axis), against Braak stage (X axis). Purple circles represents 5mC levels (OxBS), yellow circles represents 5hmC levels (BS-OxBS). Regression lines are shown for 5mC (purple line), 5hmC (yellow line) and total DNA modifications (grey line). Solid regression lines indicate P < 0.05, whilst dashed lines indicate P > 0.05*.

Figure 3.12.

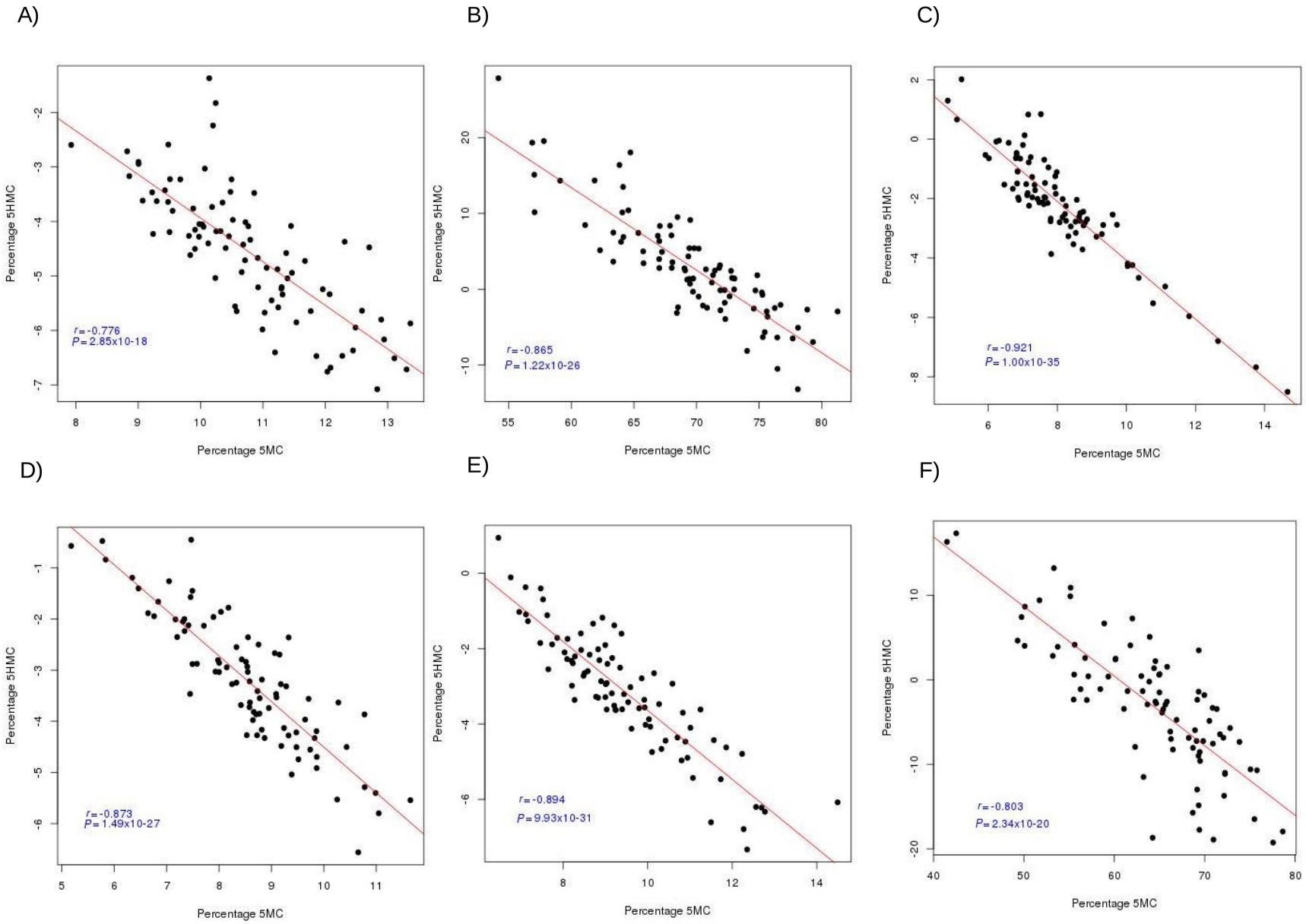
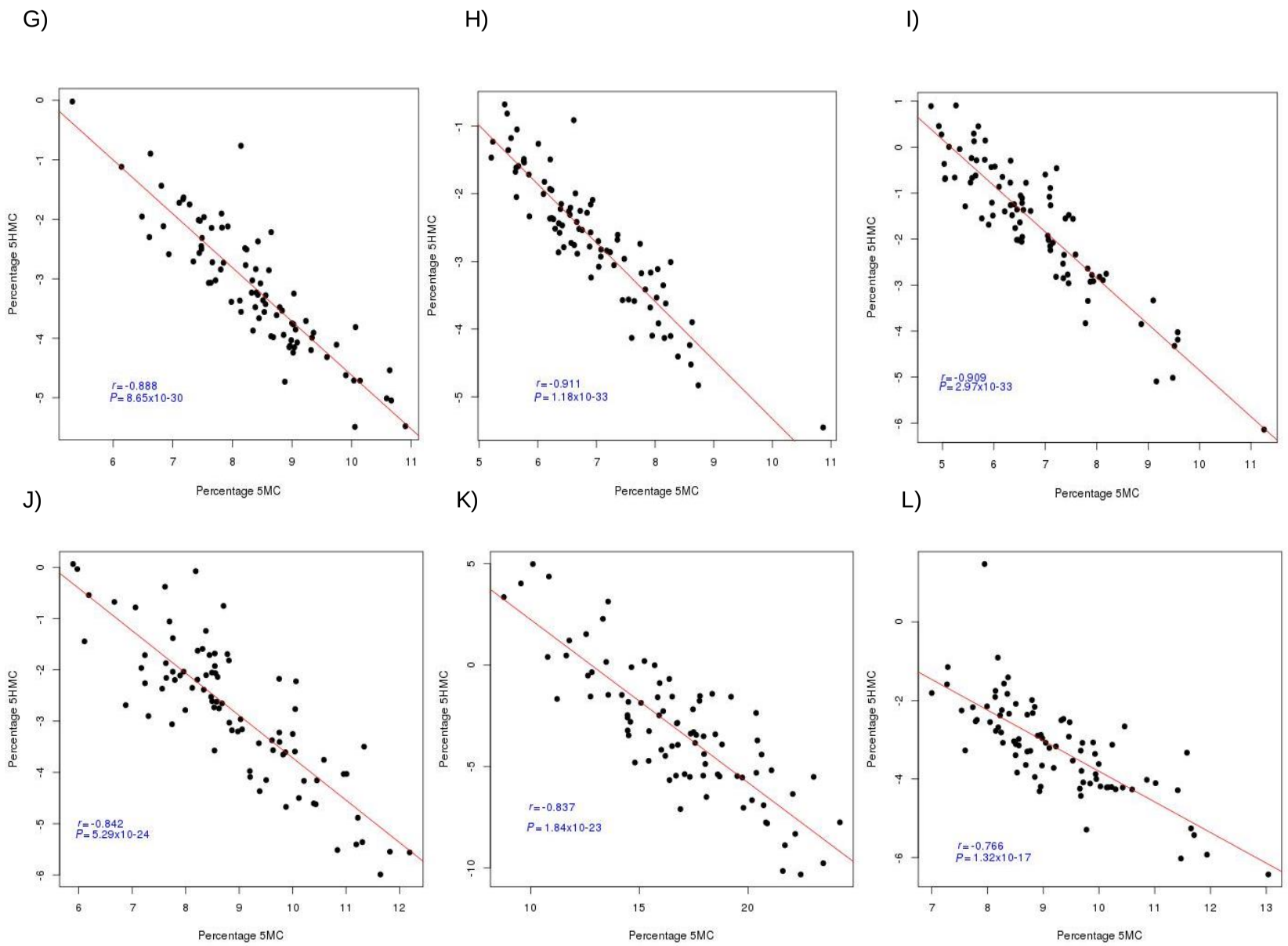

A)
B)
C)
D)
E)
F)
Percentage 5HMC
Percentage 5MC
r = -0.776
P = 2.85x10-18
r = -0.865
P = 1.22x10-26
r = -0.921
P = 1.00x10-35
r = -0.873
P = 1.49x10-27
r = -0.894
P = 9.93x10-31
r = -0.803
P = 2.34x10-20

Figure 3.12. cont.



M)

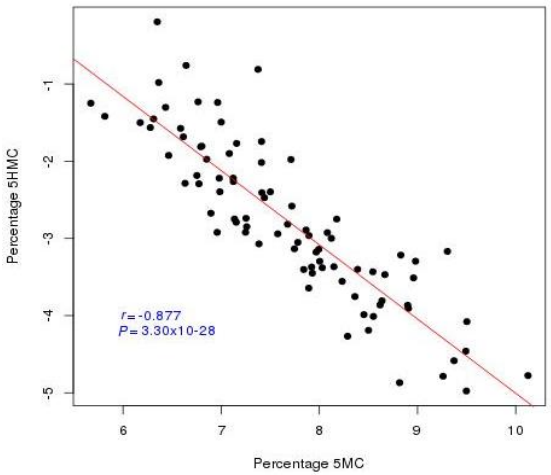


Figure 3.12; Correlation plots of 5mC levels and 5hmC levels for probes that feature in Table 3.12 with no net change in total DNA modifications. *Shown are beta levels, corrected for age and sex for (A) cg15624314 (H3F3A), (B) cg08240832 (OPN3), (C) cg13876163 (TMEM53), (D) cg18070033 (NHEJ1), (E) cg17026303 (ACVR2B), (F) cg16933762 (ROBO2), (G) cg15147060 (CGGBP1), (H) cg05052335 (RAPGEF6), (I) cg03503087 (GFRA1), (J) cg11530914 (FHOD1), (K) cg05341384 (FOXF1), (L) cg07919466 (KCNC3) and (M) cg12144476 (APP). Plotted is the % 5hmC (Y axis) against % 5mC after correcting for age and sex for all samples.*

3.4.7 *ANK1* is characterised by differential DNA methylation in the AD EC

Next, we performed an analysis to identify spatially correlated regions consisting of >2 neighboring DMPs (DMRs) or DHPs (DHRs) within a 500bp sliding window (*comb-p* (Pedersen et al. 2012)). We did not identify any DHRs, but we did identify four DMRs that reached a Sidak-corrected *P* value < 0.05 (Table 3.13). This included an eight probe DMR within the *WNT5A* gene (Sidak-corrected $P = 3.17 \times 10^{-4}$), a three probe DMR within the *TRAF3* gene (Sidak-corrected *P* = 0.022), a two probe DMR within the *ANK1* gene (Sidak-corrected *P* = 0.025) and a six probe DMR within the *ARID5B* gene (Sidak-corrected *P* = 0.029). The two DMPs we identified in *ANK1* in our regional analysis have previously been nominated as DMoPs in AD EC in two independent EWAS using BS-treated DNA (De Jager et al. 2014, Lunnon et al. 2014). In the current study, we again identified a significant increase in total DNA modifications in AD at both these two sites in the EC (Figure 3.13A: cg11823178: Δ = 6.69, $P = 4.68 \times 10^{-4}$ and Figure 3.13B: cg05066959: Δ = 10.98, $P = 8.37 \times 10^{-6}$). However, this is the first study to demonstrate that these are specifically attributable to increased 5mC (cg11823178: Δ = 9.31, $P = 2.61 \times 10^{-5}$; cg05066959: Δ = 13.71, $P = 8.03 \times 10^{-5}$), with a small, but non-significant decrease in 5hmC at both sites (cg11823178: Δ = -3.20, *P* = 0.068; cg05066959: Δ = -2.72, *P* = 0.349).

Position	Gene Name	Probes in DMR	*Comb-p P* Value	*Comb-p* Sidak corrected *P* Value
chr3:55517806-55518442	*WNT5A*	8	2.53×10^{-9}	3.17×10^{-4}
chr14:103367306-103367592	*TRAF3*	3	8.12×10^{-8}	0.022
chr8:41519308-41519400	*ANK1*	2	2.89×10^{-8}	0.025
chr10:63809073-63809171	*ARID5B*	6	3.60×10^{-8}	0.029

Table 3.13; Identification of multiple-probe DMRs associated with Braak stage. *Spatially correlated P values in a 500bp sliding window were identified using comb-p (Pedersen et al. 2012). Shown are those with a Sidak-corrected P value > 0.05.*

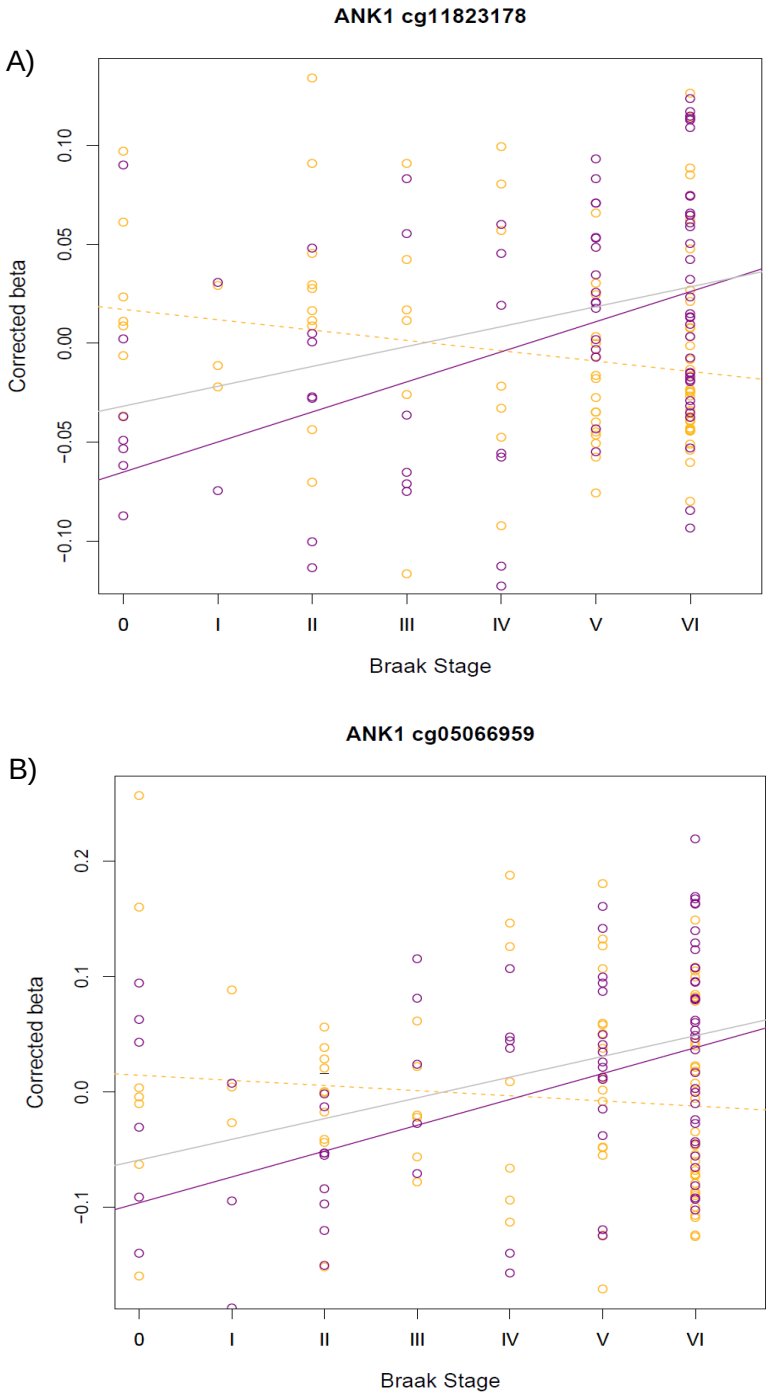


Figure 3.13; Alterations in total DNA modifications, DNA methylation and DNA hydroxymethylation are associated with AD neuropathology in the EC. *Scatter plot of beta values, corrected for age and sex, against Braak stage for probes identified in ANK1. Purple circles represent 5mC levels (OxBS), yellow circles represent 5hmC levels (BS-OxBS). Regression lines are shown for 5mC (purple line), 5hmC (yellow line) and total DNA modifications (grey line). Solid regression lines indicate P <0.05, whilst dashed lines indicate P > 0.05. Shown are (A) cg11823178 (linear regression 5mC: t_{83} = 4.2, r = 0.440, P = 2.61 x 10^{-5}; 5hmC: t_{83} = -2.0, r = -0.199, P = 0.068; total modifications: t_{89} = 3.0, r = 0.356, P = 4.68 x 10^{-4}) and (B) cg05066959 (5mC: t_{83} = 4.4, r = 0.415, P = 8.03 x 10^{-5}; 5hmC: t_{83} = -1.0, r = -0.103, P = 0.349; total modifications: t_{89} = 4.2, r = 0.440, P = 8.37 x 10^{-6}).*

3.4.8 Oxidative-bisulfite pyrosequencing validation of *ANK1*

To further confirm that *ANK1* epigenetic changes associated with neuropathology are driven by DNA hypermethylation, and not hyperhydroxymethylation, we used OxBS pyrosequencing to quantify DNA modifications across an extended region spanning eight CpG sites, including cg11823178 and cg05066959. DNA used for this study was from an independent collection of EC tissue (n = 96) obtained from the Oxford Brain Bank Collection (validation cohort 2) (Table 3.1) (Esiri 1993). In this cohort we did not have access to Braak stage III or IV samples, and as such analysed data using a case (Braak V-VI) /control (Braak 0-II) analysis model. Of the eight CpG sites assessed, seven were characterised by significant (P < 0.05) AD-associated hypermethylation (Figure 3.14A). Interestingly, significant AD-associated hypohydroxymethylation was also seen at four of the eight CpG sites (Figure 3.14B), with total DNA modifications showing a remarkably similar profile to previously published studies (Lunnon et al. 2014) (Figure 3.14C). When we averaged total DNA modifications, 5mC and 5hmC levels across the eight CpG sites in the 120bp region, we saw a significant increase in total modifications (Δ = 0.79, P = 1.12 x 10^{-4}) and 5mC (Δ = 0.87, P = 1.23 x 10^{-05}) and a trend towards a decrease in 5hmC, (Δ = -0.40, P = 0.058) in individuals with AD, compared to controls (Figure 3.14D). This data suggests that previous studies of *ANK1* 5mC in AD have underestimated the disease-associated changes in 5mC.

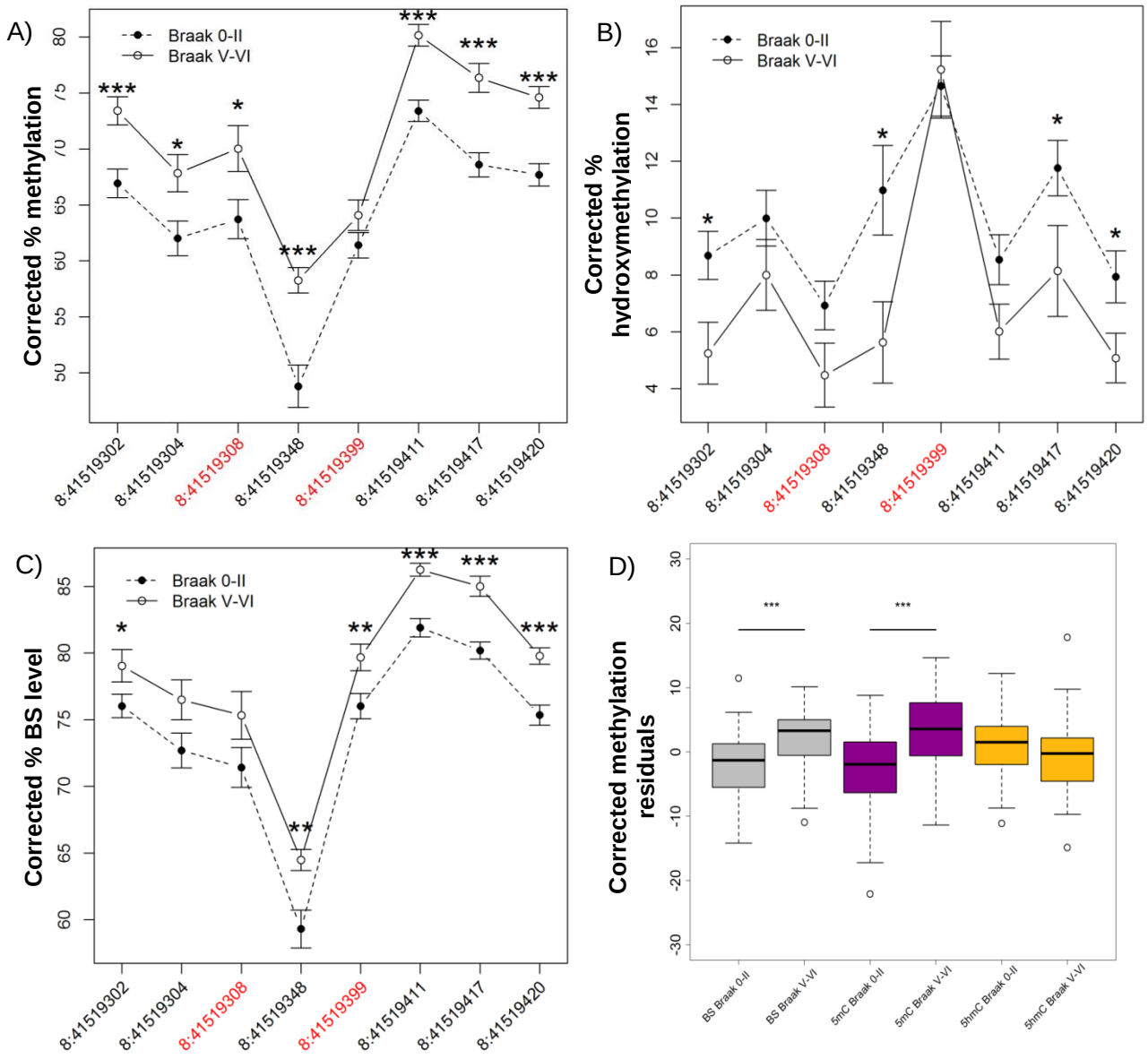

A)
Braak 0-II
Braak V-VI
Corrected % methylation
B)
Corrected % hydroxymethylation
C)
Corrected % BS level
D)
Corrected methylation residuals
8:41519302
8:41519304
8:41519308
8:41519348
8:41519399
8:41519411
8:41519417
8:41519420
BS Braak 0-II
BS Braak V-VI
5mC Braak 0-II
5mC Braak V-VI
5hmC Braak 0-II
5hmC Braak V-VI

Figure 3.14

Figure 3.14; Sites in ANK1 are characterised by significant DNA hypermethylation and hypohydroxymethylation in AD. *Using BS and OxBS pyrosequencing we determined (A) 5mC (B) 5hmC and (C) total DNA modification levels in the EC in AD samples (Braak V-VI) compared to control samples (Braak 0-II) in validation cohort 2. We assayed a 120bp region containing cg05066959 (chr8:41519308) and cg1182378 (chr8:41519399) (shown in red). We demonstrated significant (P < 0.05) neuropathology-associated hypermethylation at seven of the eight CpG sites (chr8:41519302: t_{90} = 3.7, P =3.94x10^{-4}; chr8:41519304: t_{90} = 2.6, P =0.01; chr8:41519308: t_{90} = 2.4, P =0.02; chr8:41519348: t_{90} = 4.4, P = 1.75x10^{-4}; chr8:41519411: t_{90} = 5.0, P =2.03x10^{-6}; chr8:41519417: t_{90} = 4.8, P = 6.58x10^{-6}; chr8:41519420: t_{90} = 5.1, P = 1.98x10^{-6}), hypohydroxymethylation at four of the eight CpG sites (chr8:41519302: t90 = -2.5, P =0.01; chr8:41519348: t90 = -2.4, P = 0.03; chr8:41519417: t90 = -2.0, P = 0.05; chr8:41519420: t90 = -2.3, P = 0.02), and increased total DNA modifications at six of the eight sites (chr8:41519302: t_{94} = 2.6, P =0.01; chr8:41519348: t_{94} = 3.4, P = 2.32x10^{-3}; chr8:41519399: t_{94} = 2.4, P = 0.02; chr8:41519411: t_{94} = 4.9, P = 4.83x10^{-6}; chr8:41519417: t_{94} = 4.7, P =1.03x10^{-5}; chr8:41519420: t_{94} = 4.2, P =7.90x10^{-5}). P values were calculated using linear regression; data is represented as mean (±SEM). (D) Combined analysis of averaged sites within the 120bp amplicon region highlighted a significant increase in total modifications (t_{94} = 4.0, P = 1.11 x 10^{-4}) and 5mC (t_{90} = 4.6, P = 1.23 x 10^{-05}), meeting our Bonferroni significance threshold P = 0.017 and a trend towards a decrease in 5hmC (t_{90} = -1.9, P = 0.058) in individuals with AD, compared to controls. Box and whisker plot shows median, upper and lower quartiles, and maximum and minimum values (excluding outliers) P values were calculated using two sample t-test. Key: *P < 0.05, **P < 0.01, ***P < 0.005.*

3.5 Discussion

This study represents, to our knowledge, the first large scale cross-tissue study of 5mC and 5hmC, at base-pair level resolution in AD, using matched DNA from both affected and unaffected brain regions, utilising independent study cohorts and two independent technologies (OxBS-Illumina 450K array and OxBS-pyrosequencing). Two previous studies of 5hmC in AD have been recently published. The first used a selective chemical labelling technique to enrich for 5hmC, and then sequenced the captured libraries from a small cohort of 30 individuals with AD, MCI or no dementia (Zhao et al. 2017). The study highlighted a number of AD-associated DHPs and demonstrated that gene body 5hmC was positively correlated with cis-acting gene expression. One limitation of that study was that the analysis was unable to discriminate between 5hmC and 5mC because of the low sequencing resolution. The second study, by Ellison and colleagues, utilised reduced representation hydroxymethylation profiling (RRHP) to analyse 5hmC levels in AD hippocampus (Ellison et al. 2017). Although this study assessed 5hmC at >2 million sites, the study was limited to just three AD cases and two age-matched control subjects.

The majority of studies investigating 5hmC in AD have used immuno-histochemical based methods to quantify global levels (Chouliaras et al. 2013, Condliffe et al. 2014, Coppieters et al. 2014, Lashley et al. 2015, Celarain et al. 2016). One study reported a global decrease in 5hmC in the EC and CER in AD, but no change in 5mC (Condliffe et al. 2014). However, other studies have reported global reductions in both modifications in the hippocampus (Chouliaras et al. 2013) and another study highlighted increased levels of 5mC and 5hmC in the middle frontal gyrus and middle temporal gyrus (Coppieters et al. 2014). Our results showed a net increase in 5mC and decrease in 5hmC in the EC of AD post-mortem brain, confirming previous reports within this tissue.

With a sample number of N = 85 at a set significance threshold of $P < 5 \times 10^{-5}$ we have 81.3% power to detect a moderate effect of $r = 0.5$ (Figure 3.15). In addition, in our study the overlap between DMoPs identified across our sample cohorts and those identified by Lunnon et al. (2014), despite the use of independent study

cohorts, suggests that our study was adequately powered to detect robust AD-associated differences that can be replicated in other studies. This also suggests that although different BS conversion methods and donors were used, there are consistent and reproducible changes in total DNA modifications in the EC of donors with AD.

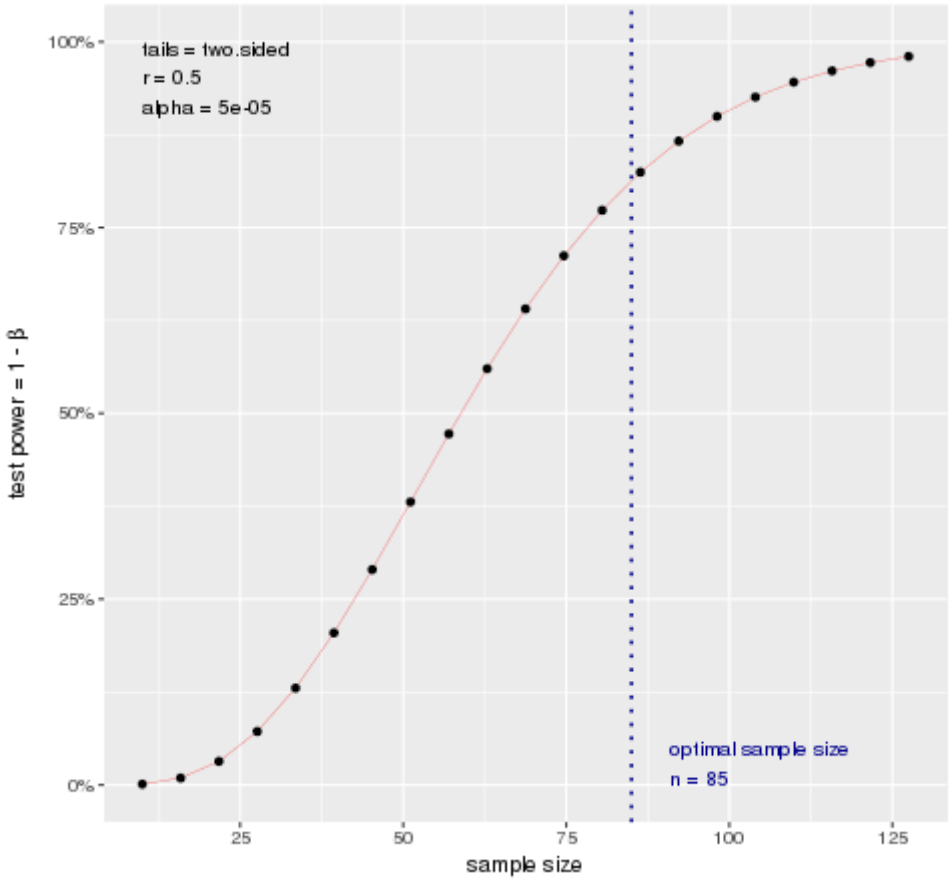


Figure 3.15; Correlation Power Calculation. *With a sample number of N = 85, there is a 81.3% power to detect r = 0.5.(Champely et al. 2018)*

We identified three Braak-associated DMPs that reached experiment-wide significance (*CGGBP1 WNT5B GNL3*) *CCGBP1* is known to influence the expression of the *FMR1* gene, which is associated with the fragile X mental retardation syndrome (Deissler et al. 1997), whereas *GNL3* is hypothesised to interact with p53, a tumor suppressor protein (Tsai 2015). *WNT5B* is part of the wnt signaling pathway, previously associated with haematopoiesis (Van Den Berg et al. 1998). However, disruption of this pathway has been previously implicated in neurodegeneration (Caricasole et al. 2004, Toledo et al. 2010).

Our pathway analysis of genes annotated to DMPs and DHPs identified a number of significant relevant pathways of interest, including several synapse related pathways that were significantly enriched for DHPs. This suggests that synaptic activity is a major pathway that displays altered levels of 5hmC in AD. Interestingly, previous work by Khare et al. (2012) has shown that 5-hmC is enriched in genes with synapse-related functions in both human and mouse brain. We have also demonstrated enrichment for the most significant DMPs and DHPs across all gene features, most strikingly Braak-associated hypomethylation and hyperhydroxymethylation appears to be enriched in the CGIs of proximal promotors. This shows that there could be a change from 5mC to 5hmC in these regions in AD, despite a global decrease (across regions covered by the array) in 5hmC with disease.

We identified 14 sites (annotated to 14 genes) that showed significant changes across more than one analysis (total modifications, 5mC alone and 5hmC alone), 13 of which showed no difference in total modifications. This is consistent with findings from previous EWAS using BS DNA, which had also not shown these loci to be associated with AD-neuropathology. One of the genes identified was *APP*, mutations in this gene are associated with early-onset, familial AD (Scheuner et al. 1996). However, our study is the first time epigenetic variation in this gene has been associated with AD via EWAS, although work by West et al. (1995) has previously identified hypomethylation of *APP* in AD brain.

We did not identify any DHRs; however, given that 5hmC is known to be enriched in areas of low CG density (Lunnon et al. 2016), this is perhaps not surprising. We did

identify four DMRs within four genes (*WNT5A, TRAF3, ANK1* and *ARID5B*). The two DMPs we identified in *ANK1* in our regional analysis have previously been highlighted as DMoPs in AD EC in two independent EWAS using BS-treated DNA (De Jager et al. 2014, Lunnon et al. 2014). We replicated the significant increase in total DNA modifications in AD at both these two sites in the EC, and demonstrated that previous studies of 5mC in *ANK1* in AD have underestimated increments in 5mC in disease due to confounding by decreased 5hmC.

Several recent studies have reported the links between epigenetic alterations at the *ANK1* locus and the development of AD (De Jager et al. 2014, Lunnon et al. 2014, Mastroeni et al. 2017). The *ANK1* protein is a plasma-bound membrane protein that contains an ankyrin repeat domain, which modulates interactions between cytoskeletal and membrane proteins (Voronin et al. 2008). One of the main functions of *ANK1* is compartmentalisation and maintenance of the plasma membrane and it is possible that the altered expression of this gene could lead to cell membrane dysfunction in AD (Lunnon et al. 2014). It has been shown that microglia, the brain's resident macrophages, have >4-fold increase in *ANK1* expression in AD compared to control brain samples (Mastroeni et al. 2017), suggesting it is possible that *ANK1* 5mC levels will have a downstream effect on immune and inflammatory regulation in the brain.

Although our study represents the first cross-tissue EWAS to measure 5mC and 5hmC in parallel in AD it is not without its own limitations. First, there are some caveats to using the 450K array, for example the presence of probes targeting regions containing SNPs and cross-hybridising probes, although these were addressed in our analysis pipeline (Section 2.8). In addition, although the 450K array is said to allow an "epigenome-wide" analysis, it only actually covers approximately 485,000 CpG positions of the 28 million CpG sites in the human genome (Lovkvist et al. 2016), similarly only a small percentage of potential non-CpG methylation sites are covered, meaning >98% of potential modification sites are not assessed. Second, although we identified DMPs that reached experiment-wide significance, the study would have benefitted from a larger discovery sample size, to allow us to identify experiment-wide significant DHPs. Third, our study was performed on "bulk" tissue; there is considerable cellular heterogeneity within brain regions and between

individuals, with the abundance of specific cell types being altered in AD. Although we used a bioinformatic algorithm (*CETS* (Guintivano et al. 2013)) to control for differences in different cell proportions, it would have been optimal to epigenetically profile different cell populations in our study. Finally, we are unable to determine whether the cytosine modification changes we have identified are the cause or consequence of disease.

3.6 Conclusions

Our analyses from multiple independent cohorts provide further evidence for a role for DNA hypermethylation, coupled with hypohydroxymethylation, across a region in *ANK1* in AD-associated neuropathology. It indicates that previous estimates of hypermethylation in *ANK1* were underestimates as it is potentially confounded by this hypohydroxymethyation. Our study identifies 13 novel loci that show no significant difference in overall DNA modifications, but display significant (opposing) differences in 5mC and 5hmC in AD. As these genes were not identified in standard BS-treated DNA, these findings highlight the utility of our method for identifying novel epigenetically dysregulated genes in AD. Although some of the AD-associated loci identified in our study do not have a mean 5hmC level across all samples >0, this likely reflects experimental noise resulting from the calculation of 5hmC from two separate arrays. However, given that the loci we identified are located within genes previously implicated in AD, such as *APP*, it is likely that our data is highlighting novel epigenetically altered loci in disease. To conclude, these novel genes warrant further research in AD.

CHAPTER 4: *ANK1* DNA MODIFICATION IN DIFFERENT NEURODEGENERATIVE DISEASES

The work presented in this chapter is based on a work submitted for publication and is currently under review at Neurobiology of Aging (Smith et al. 2018). A copy of the submitted manuscript can be found in Appendix 4.

4.1 Introduction

Dementia encompasses a group of chronic neurodegenerative diseases that affected an estimated 46.8 million people worldwide in 2013 (Prince et al. 2015), of which AD accounts for ~60% of cases. AD has been hypothesised to have an epigenetic contribution to disease aetiology (Lunnon et al. 2013). In 2014 two EWAS of AD were published, both using the 450K array and highlighting significantly increased DNA modifications in the *ANK1* gene in AD cortex (De Jager et al. 2014, Lunnon et al. 2014), which has been shown across multiple independent study cohorts (Smith et al. 2017) and in Chapter 3 of this thesis. In addition, a GWAS of a Han Chinese population published in 2015 identified a SNP in *ANK1* that conveyed an increased susceptibility for developing AD (Chi et al. 2015). ANK1 links integral membrane proteins to the underlying spectrin-actin cytoskeleton and plays key roles in activities such as cell motility, activation, proliferation, contact, and maintenance of specialised membrane domains (Yang et al. 2011); further information regarding *ANK1* can be found in Section 1.5. There is now increasing interest in understanding the role of epigenetic changes in *ANK1* in AD development and progression. One important question to be addressed relates to the specificity of *ANK1* hypermethylation to AD. Although AD accounts for ~60% of dementia cases, many other dementias share common symptoms and/or pathological hallmarks with AD.

4.1.1 Different types of dementias and neurodegenerative diseases

Although AD is the most common form of dementia, ~40% is caused by other neurodegenerative diseases (Figure 4.1). VaD and DLB explain 17% and 4% of dementia cases, respectively (Prince et al. 2014). Other dementias include HD and PD with dementia. A number of dementias have overlapping symptoms, making the diagnosis of a specific dementia difficult. In addition different dementias can often co-occur, which is called mixed dementia, for example AD and DLB, or AD and VaD. Together AD, VaD, DLB, HD, PD dementia and mixed dementia make up the vast majority (~95%) of dementia cases (Figure 4.1) (Prince et al. 2014).

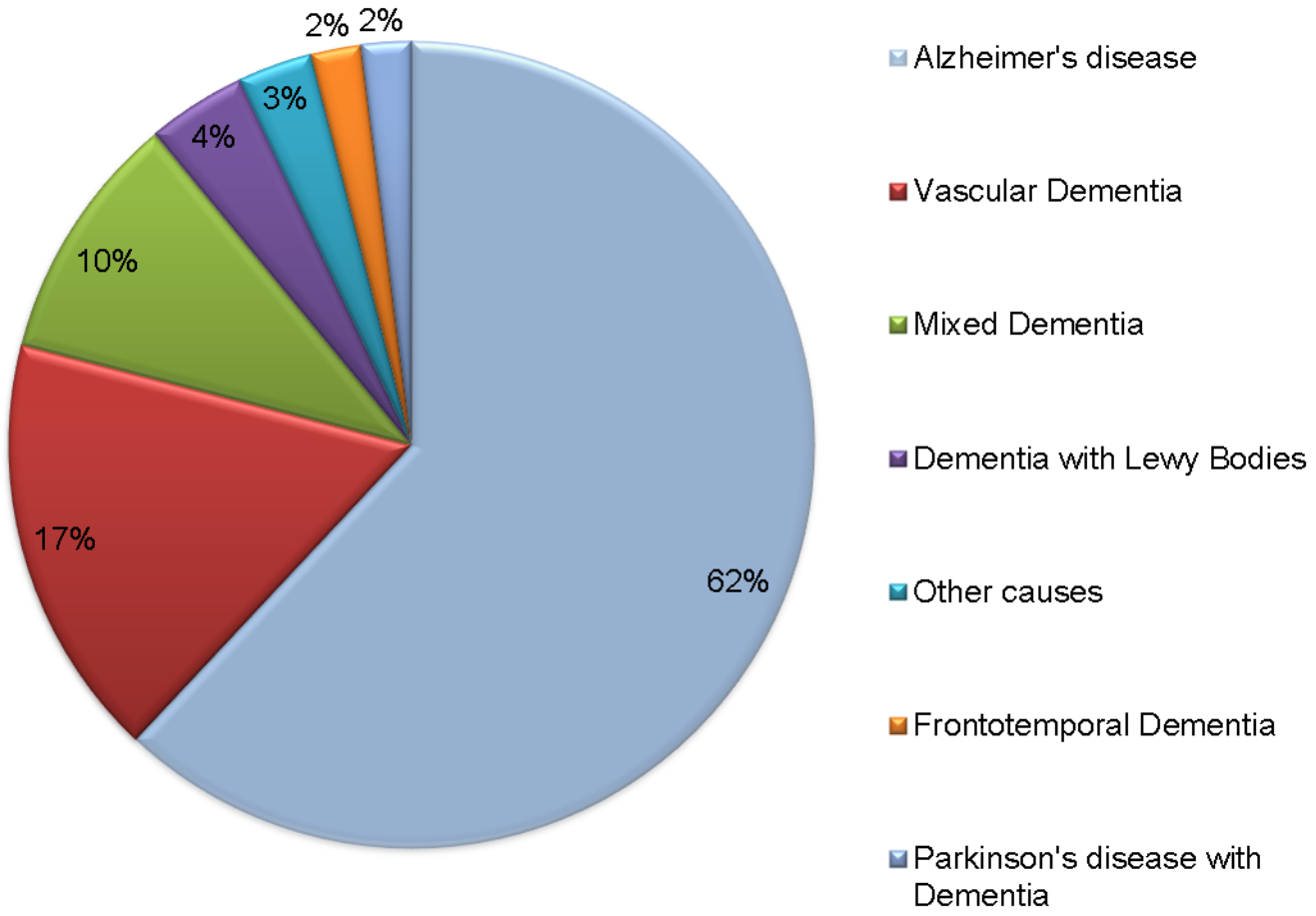


Figure 4.1; The proportion of dementia cases caused by different diseases. *Figures taken from Prince et al. (2014).*

4.1.1.1 Vascular Dementia (VaD)

VaD is the second most common cause of dementia after AD (Prince et al. 2014). Characterised by the loss of neurological function due to ischemic events, the risk of developing VaD is closely linked to vascular health (Román et al. 1993). It can be caused by a wide variety of vascular disorders, including large, or multiple small infarcts or cerebral amyloid angiopathy (CAA). VaD is most common in older people, showing a similar age demographic to AD. It is thought that blockage or damage to small vessels leads to ischaemic events, damaging neurones and secondary structures within the brain. The symptoms of VaD can occur either suddenly, following a stroke, or over time, through a series of small ischaemic events. Symptoms of VaD include; slowness of thought, difficulty with planning, trouble with understanding, problems with concentration and mood or behavioural changes (NHS 2017). Given the similarity in symptoms and age demographic for VaD and AD, it can sometimes lead to misdiagnosis of AD. In fact, it is estimated that ~17% of VaD cases are misdiagnosed (Hunter et al. 2015). Risk factors for VaD are those normally associated with cardiovascular health, such as hypertension, cigarette smoking, myocardial infarction and diabetes mellitus (Gorelick 2004). Like AD, VaD has no treatment that can halt the disease, and prescribed medications only combat vascular symptoms, for example anti-hypertensives (Rouch et al. 2015). VaD has no familial inherited cause except for one rare disorder (cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy - CADASIL) (Joutel et al. 1996). VaD can often co-occur with other dementias; it is predicted that a third of patients diagnosed with VaD will also have AD-type pathology at autopsy (Kalaria et al. 1999).

4.1.1.2 Dementia with Lewy Bodies (DLB)

DLB is the third most common cause of dementia (McKeith 2002). The pathology of DLB shares similarities to AD, with the presence of immune regulation and microglial activation being consistent between diseases (Mackenzie 2000). However, the presence of Lewy bodies within the brain makes DLB considerably more comparable to PD (McKeith 2002). DLB is characterised by the deposition of cortical and subcortical Lewy bodies, which are abnormal protein aggregates within nerve cells

(Spillantini et al. 1997). Amyloid pathology is also common in DLB however neurofibrillary tangles are rarely seen (Gomperts et al. 2008). The age of onset of this disease varies from as young as 50 to over 85 (LBDA 2016). There are three core clinical diagnostic features of DLB, including 1) cognitive fluctuation, 2) mild and spontaneous parkinsonism that typically presents as bradykinesia and rigidity and 3) visual hallucinations (Cummings 2004). Unlike AD, the deterioration in cognitive function does not occur in a linear manner; in fact one of the hallmarks of Lewy body dementia is the fluctuation of cognitive functioning. The fluctuation gives the impression that the patient is moving through disease stages in a stochastic manner; however this variation in cognitive functioning is constant within each stage of the disease (McKeith 2007). The clinical diagnosis of DLB can be difficult because of the variability and the overlap of symptoms between DLB and other related dementias, notably AD and PD dementia. In particular the diagnostic classification of either PD dementia or DLB is only disparate by the timing of symptoms; patients who develop cognitive symptoms and motor features of PD within one year are classed as having DLB. However, if a patient develops motor features first and no dementia symptoms until after a year, then PD is diagnosed (Meyniel et al. 2007, Dodel et al. 2008).

4.1.1.3 Huntington's disease (HD)

Unlike other neurodegenerative disorders, HD only has a genetic cause and there are no sporadic cases of disease. It is caused by an autosomal dominant tri-nucleotide repeat (CAG) mutation in the Huntingtin gene (*HTT*). CAG codes for the amino acid glutamine; people with fewer than 36 repeated glutamines in *HTT* produce the correctly functioning cytoplasmic protein (Walker 2007). However, a sequence of 36 or more glutamines results in an altered form of the *HTT* (Walker 2007), causing the build-up of cytosolic protein leading to neuronal cell death. Generally, the number of CAG repeats is related to how much this process is affected, and accounts for about 60% of the variation of the age of the onset of symptoms. The remaining variation has been attributed to other factors, for example, other genetic risk loci and environmental interactions (Walker 2007). 36-39 glutamine repeats gives rise to a reduced-penetrance form of the disease, with a much later onset and slower progression of symptoms (Myers 2004). Whilst with very large repeat counts, HD has full penetrance and can occur as early as 20 years of age.

However, the average age of onset is between 35 and 44 years of age (Myers 2004). HD symptoms are characterised by unwanted choreatic movements, behavioural and psychiatric disturbances and dementia. The cognitive decline in HD can be present long before the first motor symptoms appear, but it can also be very mild even in the advanced stages of the disease (Roos 2010).

4.1.1.4 Parkinson's disease (PD)

Lewy bodies, as in DLB, are the pathological hallmark of PD. They are α-synuclein immuno-reactive inclusions made up of a number of neurofilament proteins together with proteins responsible for proteolysis (Spillantini et al. 1997). Mutations in the α-synuclein gene are responsible for some familial forms of PD in which Lewy bodies are also seen (Stefanis 2012). Work by Desikan et al. (2015) has shown association of AD and PD with rs393152 within the extended MAPT region, suggesting genetic pleiotropy at this locus. Neuronal cell loss within the SN is particularly evident in PD (George et al. 2009). Post-mortem counts of neuronal numbers in this region have revealed a 50-70% loss of neurones in the SN of PD patients compared with the same region in unaffected individuals (George et al. 2009). This long-term degenerative disorder mainly affects the motor system, occurring many years before the onset of dementia (Jankovic 2008). The earliest symptoms include shaking, rigidity, slowness of movement, and difficulty with walking (Jankovic 2008), often termed "parkinsonism". However, dementia becomes common in the advanced stages of disease occurring in up to 50% of people with PD after six years (Pigott et al. 2015). The majority of PD cases are sporadic, but it is believed that both environmental and genetic factors are involved in disease aetiology (Warner et al. 2003).

4.1.2 Summary

There are many different neurodegenerative diseases that are characterised by dementia. AD is the most common form of dementia and has been hypothesised to have an epigenetic contribution to disease. Hypermethylation at the *ANK1* locus has been robustly shown in AD cortex in a number of published studies (De Jager et al. 2014, Lunnon et al. 2014) as well as in Chapter 3 of this thesis. However, to date the

extent to which altered *ANK1* 5mC is also associated with other neurodegenerative diseases is not currently known. Given that many different neurodegenerative diseases have similar symptoms and pathological hallmarks this study aimed to quantify 5mC levels across a 118bp region of *ANK1*, previously associated with AD, in a number of different neurodegenerative diseases.

4.2 Aims

- To build a resource of high quality DNA samples from multiple brain regions (EC, STG, CER, STR and SN) from post-mortem brain samples of individuals with different dementias and non- demented control donors
- To quantify 5mC across a 118bp region of *ANK1* using the BS pyrosequencing assay used in Chapter 3. To compare and contrast 5mC levels across the 118bp region in multiple neurodegenerative diseases, including AD, VaD, DLB, HD and PD.

4.3 Materials and Methods

4.3.1 Subjects and Samples

Post-mortem brain tissue was obtained from six different UK brain banks (the SWDBB, the LNDBB, the Manchester Brain Bank, the Oxford Brain Bank, the Cambridge Brain Bank and the Newcastle Brain Bank). In total, tissue was obtained from 60 AD (Braak V-VI), 119 DLB, 27 VaD, 22 HD, 36 PD and 105 elderly non-demented control subjects (Braak 0-II). A subset of DLB (N = 39) and VaD (N = 5) cases also had co-existing AD pathology. For each disease we analysed the EC, STG and CER. For HD cases we also analysed the STR as this is primarily affected in disease (Reiner et al. 2011), whilst for PD cases we analysed the STR and SN as these are regions of pathology in this disease (Fearnley et al. 1991). For control samples we analysed all five brain regions. For a small number of donors tissue was not available from all brain regions. Genomic DNA was isolated from ~100mg of each dissected brain region using a standard phenol-chloroform extraction method, and tested for degradation and purity prior to analysis as previously described (Smith et al. 2016) (Section 2.3). Demographic information for samples can be found in Table 4.1.

		Brain Region				
		EC	**STG**	**CER**	**STR**	**SN**
AD (N = 60)	**N**	58	58	58	-	-
	Sex (M/F)	27/31	26/32	26/32	-	-
	Mean age (±SD)	77.7 (9.2)	77.6 (9.2)	78.4 (8.8)	-	-
All DLB (N = 119)	**N**	91	107	112	-	-
	Sex (M/F)	72/19	84/23	88/24	-	-
	Mean age (±SD)	78.8 (7.3)	78.3 (7.2)	78.1 (7.2)	-	-
DLB only (N = 80)	**N**	56	69	76	-	-
	Sex (M/F)	47/9	56/13	61/15	-	-
	Mean age (±SD)	78.1 (7.1)	77.6 (6.9)	77.4 (7.0)	-	-
DLB + AD (N = 39)	**N**	35	38	36	-	-
	Sex (M/F)	25/10	28/10	27/9	-	-
	Mean age (±SD)	80.0 (7.5)	79.5 (7.6)	79.5 (7.7)	-	-
All VaD (N = 27)	**N**	23	27	26	-	-
	Sex (M/F)	9/14	11/16	11/15	-	-
	Mean age (±SD)	83.9 (8.4)	83.6 (7.9)	83.5 (8.1)	-	-
VaD only (N = 22)	**N**	19	22	21	-	-
	Sex (M/F)	8/11	10/12	10/11	-	-
	Mean age (±SD)	83.3 (7.8)	83.0 (7.4)	82.8 (7.5)	-	-
VaD + AD (N = 5)	**N**	4	5	5	-	-
	Sex (M/F)	1/3	1/4	1/4	-	-
	Mean age (±SD)	86.8 (12.1)	86.2 (10.6)	86.2 (10.6)	-	-
HD (N = 22)	**N**	19	20	20	19	-
	Sex (M/F)	13/6	13/7	14/6	13/6	-
	Mean age (±SD)	71.1 (5.0)	70.4 (7.1)	70.0 (7.2)	68.8 (12.8)	-
PD (N = 36)	**N**	33	35	33	5	29
	Sex (M/F)	19/14	21/14	21/12	4/1	19/10
	Mean age (±SD)	78.5 (7.4)	78.2 (7.5)	78.3 (7.5)	83.6 (3.2)	78.9 (7.6)
Control (N = 105)	**N**	99	74	98	21	31
	Sex (M/F)	53/46	40/34	48/50	9/12	11/20
	Mean age (±SD)	81.4 (9.6)	81.6 (10.6)	80.7 (10.7)	76.5 (11.7)	77.3 (11.2)

Table 4.1; Sample and demographic information for samples used in the study. *Shown for each disease cohort are the number of total donors used (N), then for each individual brain region, the number of samples available (N), sex distribution and mean age (± standard deviation (SD)).*

4.3.2 *ANK1* bisulfite pyrosequencing

Throughout this chapter we refer to 5mC, however as we utilised BS treatment, the data we report is actually the sum of both 5mC and 5hmC at the *ANK1* locus. Due to financial and logistical constraints it was not feasible to perform OxBS treatment in this study. BS pyrosequencing was used to quantify 5mC across eight individual CpG sites in the *ANK1* gene, spanning from 41519302 to 41519420 within chromosome 8 (hg19). BS conversion was performed using the Bisulfite-Gold kit (Zymo research, USA). A single amplicon (246bp) was generated using primers designed using the PyroMark Assay Design software 2.0 (Qiagen, UK) as previously described (Lunnon et al. 2014). Pyrosequencing was performed using two sequencing primers to maximise coverage across eight CpG sites. 5mC was quantified using the Pyromark Q24 system (Qiagen, UK) following the manufacturer's standard instructions and the Pyro Q24 CpG 2.0.6 software (Section 2.9).

4.3.3 Data Analysis

All computations and statistical analyses were performed using R 3.3.2 (R Development Core Team 2012). A linear regression analysis was performed, controlling for the effects of age, sex and batch effects, comparing control samples with samples affected by each neurodegenerative disease. Linear regression was used to maintain consistency with Chapter 3 and previous publication (Lunnon et al. 2014). For the VaD and DLB samples we also performed a second analysis to investigate whether co-existing AD pathology influenced the results by comparing individuals with and without co-existing AD pathology to control samples, again accounting for the effects of age, sex and batch. We then used a two sample t-test to assess the difference in the mean of examined 5mC differences averaged across the amplicon for each disease studied. Bonferroni significance threshold was calculated per test to adjust for multiple testing. However, it should be noted that Bonferroni correction is highly conservative and too stringent for this analysis due to the highly correlated nature of the brain regions analysed. Finally, we used a paired two-tailed t-test to compare adjusted 5mC differences in disease between brain regions.

4.4 Results

4.4.1 AD-associated *ANK1* DNA hypermethylation is seen across all tissues analysed

First, we sought to replicate previous findings of *ANK1* DNA hypermethylation in AD. Across the 118bp region, we observed significantly increased levels in AD cases compared to controls in all eight *ANK1* CpG sites in the EC (Figure 4.2A) and seven *ANK1* CpG sites in the STG (Figure 4.2B). Both the EC and STG exhibit a high degree of AD pathology, even in the earliest stages of disease, with the EC being the starting point of AD pathology in the cortex (Braak et al. 1991). Conversely, the CER remains free of AD pathology until the very last stages of the disease, although even then this is limited to β-amyloid plaques with an absence of NFTs of hyperphosphorylated tau (Braak et al. 1989). Of note, we observed significant *ANK1* DNA hypermethylation at five *ANK1* CpG sites in the CER (Figure 4.2C). This is the first time *ANK1* 5mC differences have been reported in the CER. Interestingly, two of the loci that did not display AD-associated *ANK1* hypermethylation in the CER were chr8:41519308 and chr8:41519399, the two sites included on the 450K array used in previous EWAS analyses of AD which did not identify *ANK1* hypermethylation in AD in the CER (Lunnon et al. 2014). Average 5mC across the amplicon region was significantly elevated in AD in the EC ($P = 1.29 \times 10^{-07}$), STG ($P = 2.39 \times 10^{-03}$) and CER ($P = 7.81 \times 10^{-03}$) (Figure 4.2D). *ANK1* 5mC differences between cases and controls at both individual sites and across the amplicon were lower in the CER compared to other tissues tested (Table 4.2), with a significantly greater 5mC difference between cases and controls in the EC (amplicon average corrected 5mC difference (Δ) = 4.53%) compared to both the STG (amplicon average Δ = 2.84%; $P = 7.98 \times 10^{-4}$) and the CER (amplicon average Δ = 1.17%, $P = 2.55 \times 10^{-4}$). Interestingly, this pattern of difference matches the spread of AD pathology throughout the brain.

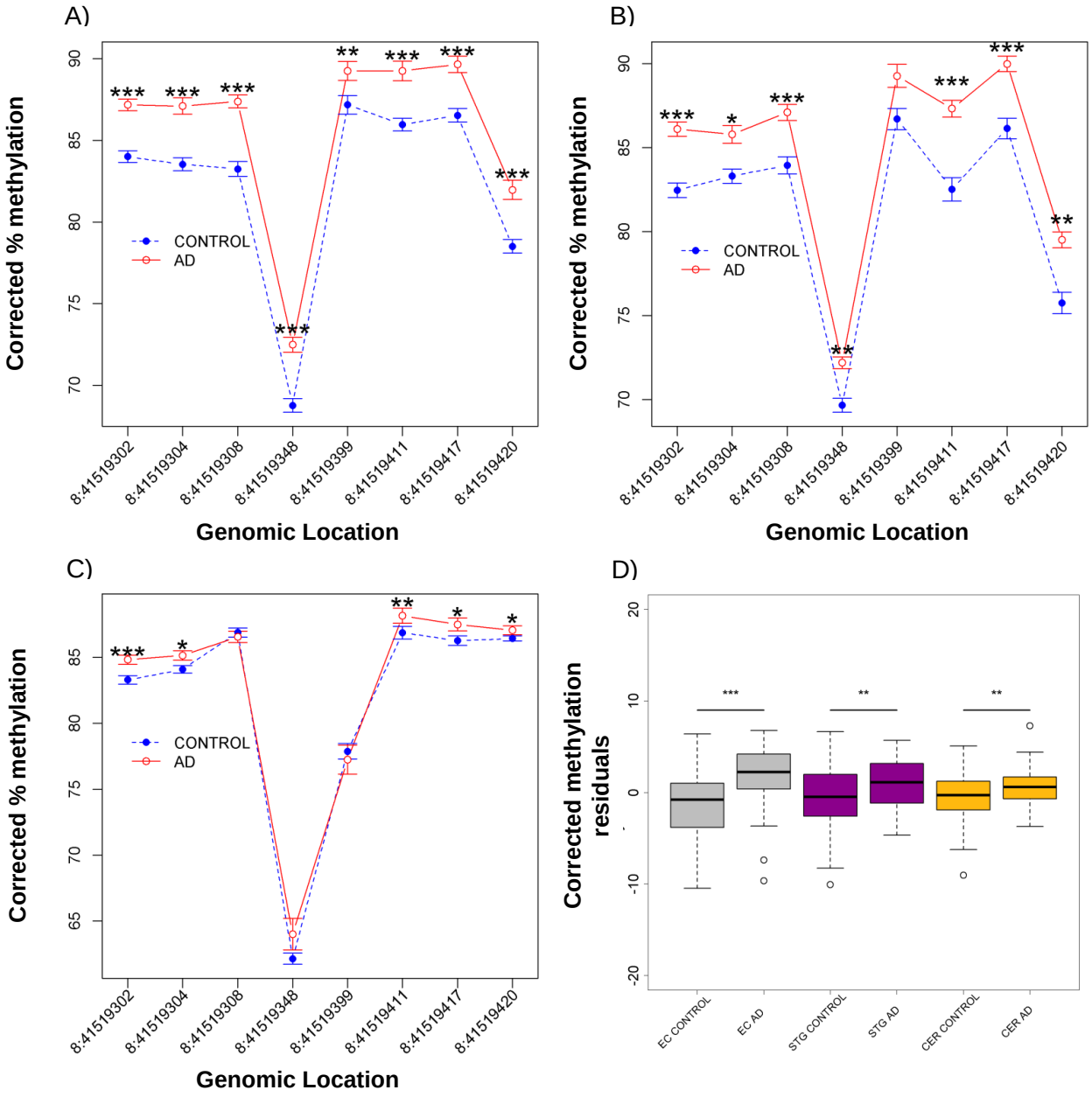

A)
B)
C)
D)
Corrected % methylation
Corrected methylation residuals
CONTROL
AD
Genomic Location
8:41519302
8:41519304
8:41519308
8:41519348
8:41519399
8:41519411
8:41519417
8:41519420
EC CONTROL
EC AD
STG CONTROL
STG AD
CER CONTROL
CER AD

Figure 4.2

Figure 4.2; ANK1 is hypermethylated in the EC, STG and CER in AD brain. *Using BS pyrosequencing we assayed a 118bp region of the ANK1 gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in (A) EC, (B) STG and (C) CER tissue, in AD samples (Braak V-VI) compared to control samples (Braak 0-II) using linear regression. We demonstrated significant neuropathology-associated hypermethylation at all assayed CpG sites in the EC (chr8:41519302: t_{153} = 6.9, r = 0.53, P =1.76x10^{-10}; chr8:41519304: t_{153} = 6.6, r = 0.51, P =5.76x10^{-10}; chr8:41519308: t_{148} = 7.2, r = 0.53, P =3.16x10^{-11}; chr8:41519348: t_{140} = 6.8, r = 0.55, P = 1.75x10^{-4}; chr8:41519399: t_{155} = 2.8, r = 0.25, P = 5.35x10^{-3}; chr8:41519411: t_{155} = 5.9, r = 0.45, P =2.27x10^{-8}; chr8:41519417: t_{155} = 5.5, r = 0.39, P = 1.50x10^{-7}; chr8:41519420: t_{155} = 5.9, r = 0.50, P = 2.53x10^{-8}), significant hypermethylation at seven of the eight CpG sites in the STG (chr8:41519302: t_{129} = 3.9, r = 0.54, P =1.49x10^{-4}; chr8:41519304: t_{129} = 2.4, r = 0.41, P =0.017; chr8:41519308: t_{121} = 3.6, r = 0.47, P =4.43x10^{-4}; chr8:41519348: t_{114} = 3.2, r = 0.50, P = 0.002; chr8:41519411: t_{129} = 3.8, r = 0.63, P =2.22x10^{-4}; chr8:41519417: t_{129} = 3.6, r = 0.44, P = 4.55x10^{-4}; chr8:41519420: t_{129} = 3.0, r = 0.45, P = 0.003) and five of the eight sites in the CER(chr8:41519302: t_{152} = 3.7, r = 0.34, P =2.98x10^{-4}; chr8:41519304: t152 = 2.7, r = 0.32, P =0.008; chr8:41519411: t_{153} = 3.4, r = 0.34, P =8.43x10^{-4}; chr8:41519417: t_{151} = 2.2, r = 0.26, P = 0.028; chr8:41519420: t_{151} = 2.1, r = 0.13, P = 0.038). Data is represented as mean (±SEM) (D) When data was summed across the 118bp amplicon region we observed a significant increase in 5mC across all brain regions (two sample t-test, EC: t_{157} = 5.5, P = 1.29 x 10^{-7}, STG: t_{132} = 3.1, P = 0.002 and CER: t_{154} = 2.8, P = 0.005). Meeting our Bonferroni significance in EC only (P = 1.85x10^{-3}). Box and whisker plot shows median, upper and lower quartiles, and maximum and minimum values (excluding outliers) Key: *P < 0.05, **P < 0.01, ***P < 0.005.*

Genomic Location	AD					
	EC		STG		CER	
	Δ	***P***	Δ	***P***	Δ	***P***
chr8:41519302	**4.31**	**1.76E-10**	**2.83**	**1.49E-04**	**2.09**	**2.98E-04**
chr8:41519304	**4.85**	**5.76E-10**	**2.01**	**0.017**	**1.42**	**0.008**
chr8:41519308	**5.51**	**3.16E-11**	**3.22**	**4.43E-04**	-0.23	0.729
chr8:41519348	**4.99**	**2.79E-10**	**2.33**	**0.002**	0.17	0.838
chr8:41519399	**2.87**	**0.005**	1.68	0.150	0.66	0.585
chr8:41519411	**4.58**	**2.27E-08**	**4.02**	**2.22E-04**	**2.83**	**8.43E-04**
chr8:41519417	**4.30**	**1.50E-07**	**3.53**	**4.55E-04**	**1.54**	**0.028**
chr8:41519420	**4.80**	**2.53E-08**	**3.08**	**0.003**	**0.86**	**0.038**
Amplicon Average	**3.14**	**1.29E-07**	**1.68**	**0.002**	**0.98**	**0.005**

Table 4.2; ANK1 CpG sites are hypermethylated in AD across all brain regions tested. *Shown for each tissue is chromosomal location (genome build hg19) of CpG sites tested, with corrected difference (Δ) in 5mC levels and corresponding P value between controls (Braak 0-II) and AD cases (Braak V-VI) after adjusting for the covariates of age, sex and batch effects using linear regression. A positive Δ reflects hypermethylation in AD. All P values < 0.05 (and their corresponding Δ) are shown in bold.*

4.4.2 *ANK1* DNA hypermethylation in the EC is only observed in DLB cases with co-existing AD pathology

Interestingly we observed significant hypermethylation of *ANK1* in DLB cases compared to controls in the EC (Figure 4.3A) at four of the eight *ANK1* CpG sites (Table 4.3). We saw no difference between DLB and control samples in either the STG (Figure 4.4A) or the CER (Figure 4.4B) at any of the eight *ANK1* CpG sites. Across the *ANK1* amplicon we observed significant DLB-associated hypermethylation in the EC (P = 0.0244), but not in the STG or CER (P > 0.05) (Figure 4.3B). It is widely reported that DLB and AD frequently co-occur (Rosenberg et al. 2001); we were therefore interested to investigate whether we still observed DLB-associated DNA hypermethylation in the EC when we stratified for a co-existing AD pathology. We found no significant difference in *ANK1* 5mC in individuals with "pure" DLB compared to controls in the EC (N = 56, Figure 4.3C), STG (N = 69, Figure 4.4C) and CER (N = 76, Figure 4.4D). However, we did observe significant hypermethylation in DLB cases with a co-existing AD pathology compared to controls at seven of the eight *ANK1* CpG sites in the EC (N = 35, Figure 4.3C) and two sites in the STG (N = 38, Figure 4.4C), with no difference in the CER (N = 36, Figure 4.4D). When we looked across the whole 118bp region, we saw increased *ANK1* 5mC in the EC in individuals with co-existing AD pathology ($P = 1.45 \times 10^{-03}$) (Figure 4.3D), suggesting that the *ANK1* hypermethylation seen in some individuals with DLB is primarily driven by AD pathology.

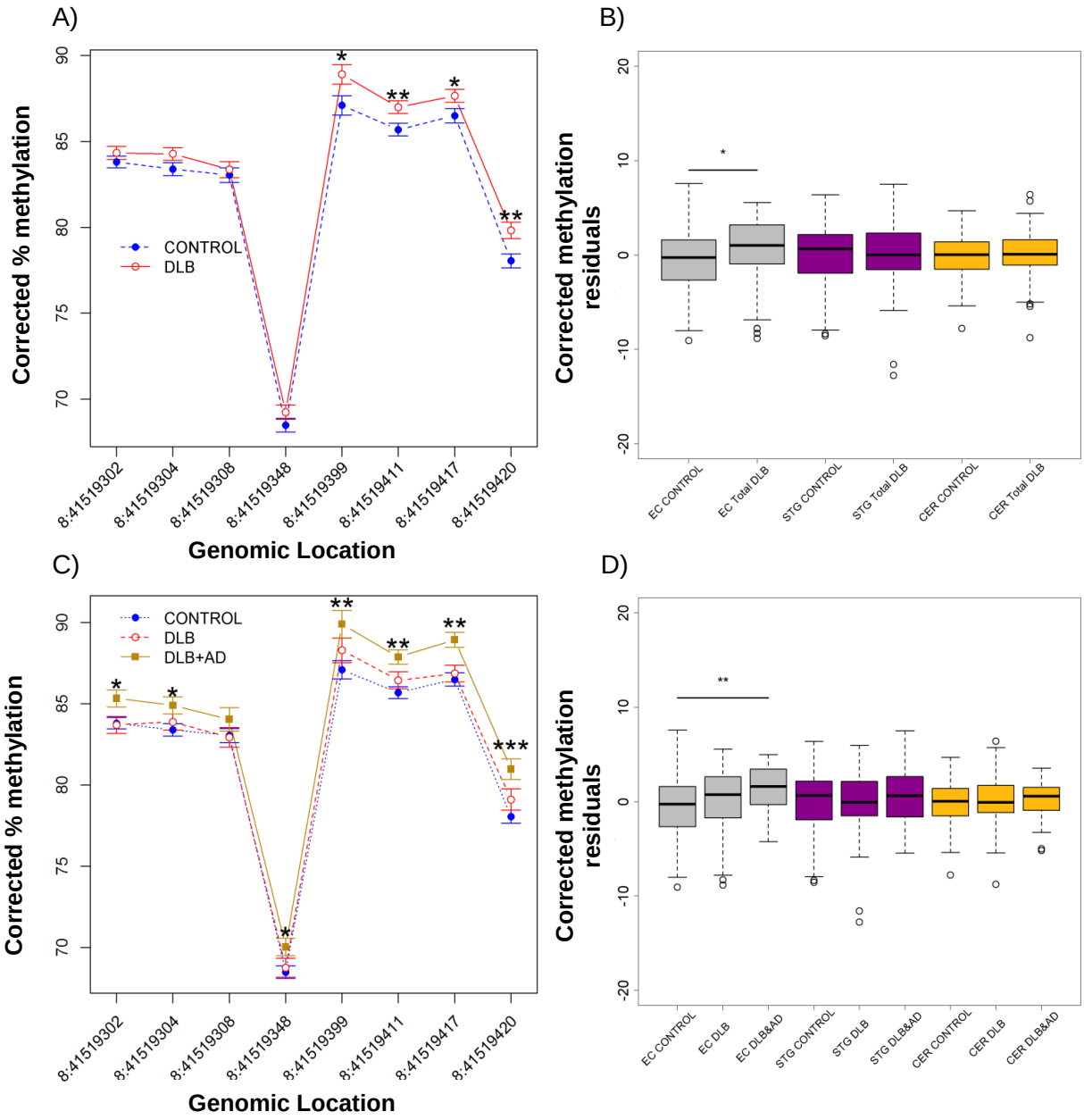

A)
Corrected % methylation
CONTROL
DLB
8:41519302
8:41519304
8:41519308
8:41519348
8:41519399
8:41519411
8:41519417
8:41519420
Genomic Location
B)
Corrected methylation residuals
EC CONTROL
EC Total DLB
STG CONTROL
STG Total DLB
CER CONTROL
CER Total DLB
C)
Corrected % methylation
CONTROL
DLB
DLB+AD
Genomic Location
D)
Corrected methylation residuals
EC CONTROL
EC DLB
EC DLB&AD
STG CONTROL
STG DLB
STG DLB&AD
CER CONTROL
CER DLB
CER DLB&AD

Figure 4.3

Figure 4.3; ANK1 hypermethylation is observed in the EC in individuals with DLB and co-existing AD pathology. *Using BS pyrosequencing we assayed a 118bp region of the ANK1 gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in all DLB samples compared to control samples. (A) We demonstrated significant neuropathology-associated hypermethylation at four CpG sites in the EC. (linear regression: chr8:41519399: t_{185} = 2.4, r = 0.24, P = 0.020; chr8:41519411: t_{185} = 2.6, r = 0.28, P =0.010; chr8:41519417: t_{185} = 2.2, r = 0.24, P = 0.032; chr8:41519420: t_{185} = 3.0, r = 0.31, P = 0.003). (B) When data was summed across the 118bp amplicon region we observed a significant increase in 5mC in the EC (two sample t-test, t_{188} = 2.3, P = 0.024), which did not meet our stringent Bonferroni significance threshold ($1.85x10^{-3}$). Some individuals with DLB also had co-existing AD pathology; (C) when we compared 5mC levels in ANK1 in individuals with DLB and co-existing AD pathology, or individuals with "pure" DLB only to controls, we found significant hypermethylation at seven sites in the EC in individuals with co-existing AD pathology(chr8:41519302: t_{131} = 2.3, r = 0.63, P =0.022; chr8:41519304: t_{131} = 2.1, r = 0.18, P =0.034; chr8:41519348: t_{122} = 2.1, r = 0.17, P = 0.034; chr8:41519399: t_{132} = 2.6, r = 0.21, P = 0.010; chr8:41519411: t_{132} = 3.3, r = 0.26, P =0.002; chr8:41519417: t_{132} = 3.3, r = 0.26, P = 0.001; chr8:41519420: t_{132} = 3.7, r = 0.30, P = $2.45x10^{-4}$). (D) When we averaged methylation across the region we saw significant hypermethylation in the EC in individuals with co-existing AD pathology (two sample t-test, t_{132} = 3.3, P = 0.001), meeting our stringent Bonferroni significance threshold ($1.85x10^{-3}$). Data is represented as mean (±SEM) Box and whisker plots show median, upper and lower quartiles, and maximum and minimum values (excluding outliers). Key: *P < 0.05, **P < 0.01, ***P < 0.005.*

	All DLB						DLB Only						DLB + AD					
	EC		STG		CER		EC		STG		CER		EC		STG		CER	
Genomic Location	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***
chr8:41519302	0.58	0.286	0.34	0.516	0.24	0.542	0.10	0.863	0.02	0.963	0.16	0.712	**1.53**	**0.022**	0.81	0.198	0.34	0.552
chr8:41519304	0.97	0.083	0.74	0.196	0.04	0.907	0.49	0.436	0.26	0.669	0.05	0.892	**1.51**	**0.034**	**1.42**	**0.034**	0.01	0.986
chr8:41519308	0.35	0.600	0.54	0.368	0.42	0.419	0.12	0.872	0.08	0.895	0.58	0.261	1.00	0.230	1.20	0.110	0.06	0.938
chr8:41519348	0.84	0.171	0.89	0.114	0.35	0.580	0.27	0.692	0.39	0.492	0.46	0.484	**1.55**	**0.034**	**1.58**	**0.025**	-0.02	0.985
chr8:41519399	**1.99**	**0.020**	0.68	0.388	0.31	0.743	1.19	0.208	0.70	0.399	0.25	0.791	**2.81**	**0.010**	0.45	0.645	0.31	0.791
chr8:41519411	**1.44**	**0.010**	0.03	0.970	0.12	0.829	0.76	0.228	0.17	0.845	-0.01	0.993	**2.19**	**0.001**	0.39	0.716	0.36	0.664
chr8:41519417	**1.28**	**0.032**	-0.22	0.758	0.03	0.946	0.37	0.588	0.55	0.468	-0.07	0.888	**2.44**	**0.001**	0.44	0.635	0.25	0.696
chr8:41519420	**1.95**	**0.003**	-0.75	0.302	-0.29	0.381	1.05	0.150	0.70	0.375	-0.41	0.272	**2.92**	**2.45E-04**	0.64	0.502	0.05	0.888
Amplicon Average	**1.08**	**0.024**	0.23	0.641	0.16	0.623	0.51	0.370	0.06	0.920	0.14	0.706	**1.99**	**0.001**	0.75	0.229	0.20	0.638

Table 4.3; ANK1 CpG sites are hypermethylated in the EC in patients with DLB and co-existing AD pathology. *Shown for each tissue is chromosomal location (genome build hg19) of CpG sites tested, with corrected difference (Δ) in 5mC levels and corresponding P value between controls and DLB cases after adjusting for the covariates of age, sex and batch effects (linear regression). Analyses are stratified into "All DLB", "DLB Only" and "DLB + AD" groups. A positive Δ reflects hypermethylation in disease. All P values < 0.05 (and their corresponding Δ) are shown in bold.*

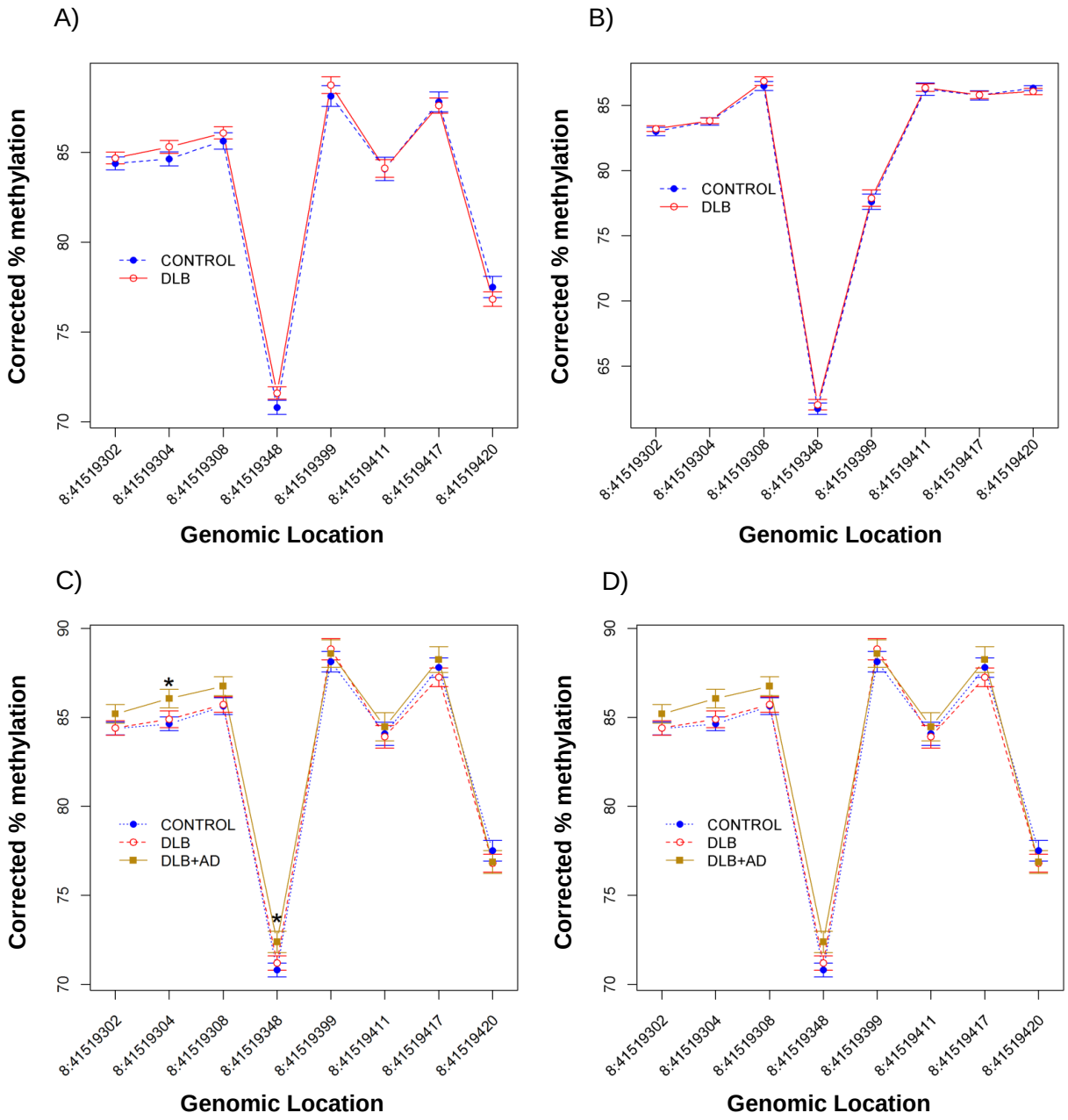


Figure 4.4; ANK1 DNA methylation levels in the STG and CER in individuals with DLB *Using BS pyrosequencing we assayed a 118bp region of the ANK1 gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in all DLB samples compared to control samples. We did not observe any disease-associated differences in either the STG (A) or CER (B). When we compared 5mC levels in ANK1 in individuals with DLB and co-existing AD pathology, or individuals with "pure" DLB only to controls, we found significant hypermethylation at two sites in the STG (linear regression, chr8:41519304: t109 = 2.1, r = 0.18, P =0.034; chr8:41519348: t101 = 2.3, r = 0.20, P = 0.025) (C) but no sites in the CER (D) in individuals with co-existing AD pathology. Data is represented as mean (±SEM) Key: *P < 0.05, **P < 0.01, ***P < 0.005.*

4.4.3 *ANK1* hypermethylation is seen in the EC only in VaD individuals with co-existing AD pathology

We observed increased 5mC in individuals with VaD at none of the *ANK1* CpG sites in the EC (Figure 4.5A), only one site in the STG (Figure 4.6A) and no sites in the CER (Figure 4.6B) (Table 4.4), with no difference across the amplicon in any of the brain regions tested ($P > 0.05$) (Figure 4.5B). Because VaD also often co-occurs with AD we next examined whether stratifying cases by the presence of AD pathology altered these findings. Interestingly, we saw disease-associated hypermethylation in the EC at five of the eight ANK1 CpG sites only in individuals with co-existing AD pathology (Figure 4.5C), whilst we saw disease-associated hypomethylation at one site in the STG (Figure 4.6C) in individuals with “pure” VaD, and no disease-associated changes in the CER (Figure 4.6D). When we looked across the 118bp region we only saw significant *ANK1* hypermethylation in individuals with VaD and co-existing AD pathology compared to controls in the EC ($P = 0.0163$) (Figure 4.5D). It is worth noting that our cohort only had a small number of VaD cases with co-existing AD pathology (N=5).

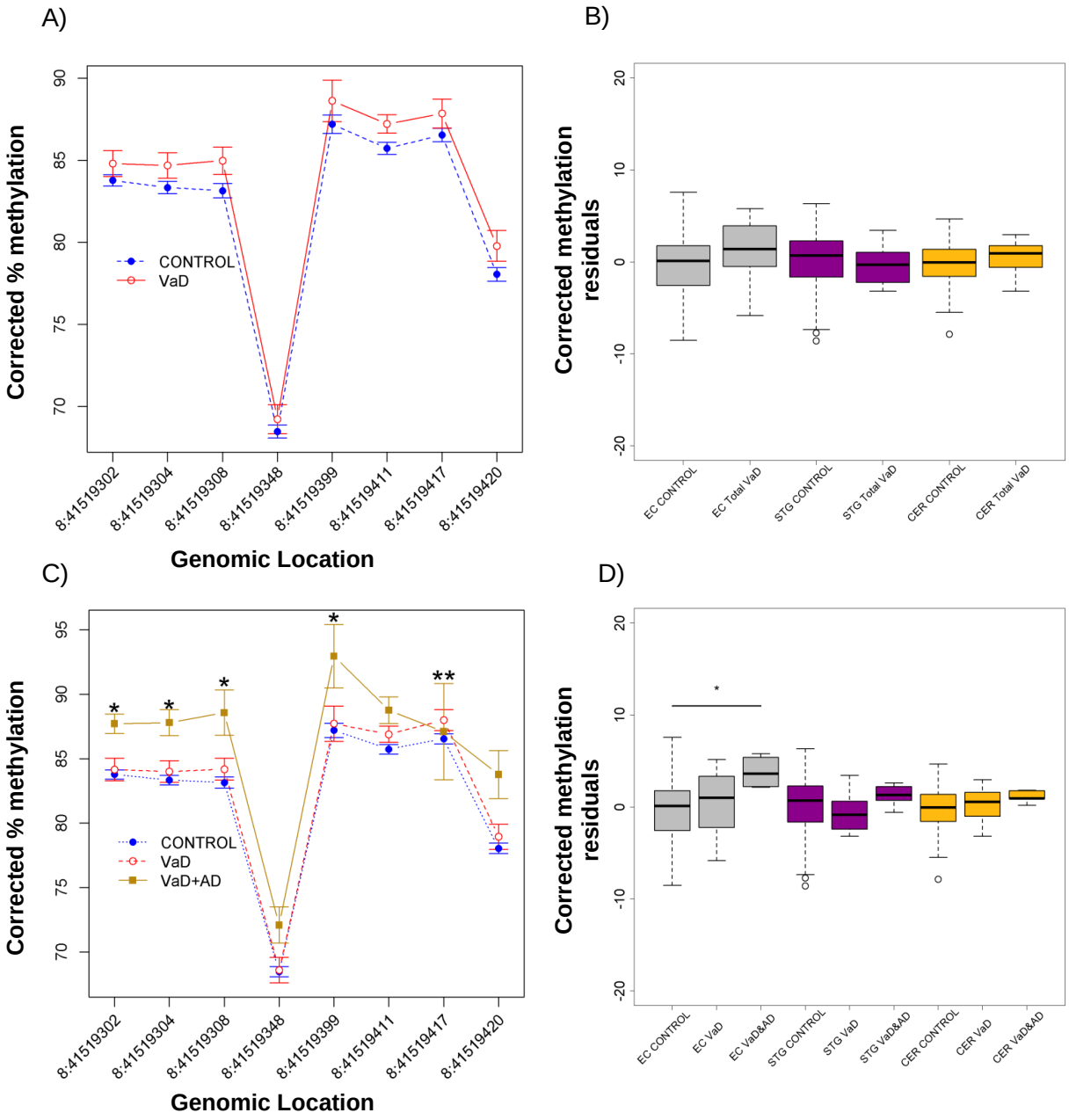

A)
B)
C)
D)
Corrected % methylation
Genomic Location
CONTROL
VaD
VaD+AD
Corrected methylation residuals
8:41519302
8:41519304
8:41519308
8:41519348
8:41519399
8:41519411
8:41519417
8:41519420
EC CONTROL
EC Total VaD
STG CONTROL
STG Total VaD
CER CONTROL
CER Total VaD
EC VaD
EC VaD&AD
STG VaD
STG VaD&AD
CER VaD
CER VaD&AD

Figure 4.5

Figure 4.5; ANK1 hypermethylation is observed in the EC in individuals with VaD and co-existing AD pathology. *Using BS pyrosequencing we assayed a 118bp region of the ANK1 gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in all VaD samples compared to control samples. (A) We found no disease-associated differential methylation at any individual sites in the EC, or any difference when averaged across the amplicon (B). Some individuals with VaD also had co-existing AD pathology; (C) when we compared 5mC levels in ANK1 in individuals with VaD and co-existing AD pathology, or individuals with "pure" VaD only to controls, we found significant hypermethylation at five sites in the EC in individuals with co-existing AD pathology. (linear regression: chr8:41519302: $t_{100} = 2.3$, $r = 0.20$, $P = 0.027$; chr8:41519304: $t_{100} = 2.4$, $r = 0.21$, $P = 0.020$; chr8:41519308: $t_{98} = 2.6$, $r = 0.23$, $P = 0.012$; chr8:41519399: $t_{101} = 2.0$, $r = 0.17$, $P = 0.045$; chr8:41519417: $t_{101} = 2.7$, $r = 0.24$, $P = 0.008$) (D) When we averaged methylation across the region we observed significant hypermethylation in the EC in individuals with a co-diagnosis of AD (two sample t-test: $t_{101} = 2.4$, $P = 0.016$), which did not meet our stringent Bonferroni significance threshold (1.85×10^{-3}). Data is represented as mean (±SEM). Box and whisker plots show median, upper and lower quartiles, and maximum and minimum values (excluding outliers). Key: *P < 0.05, **P < 0.01, ***P < 0.005.*

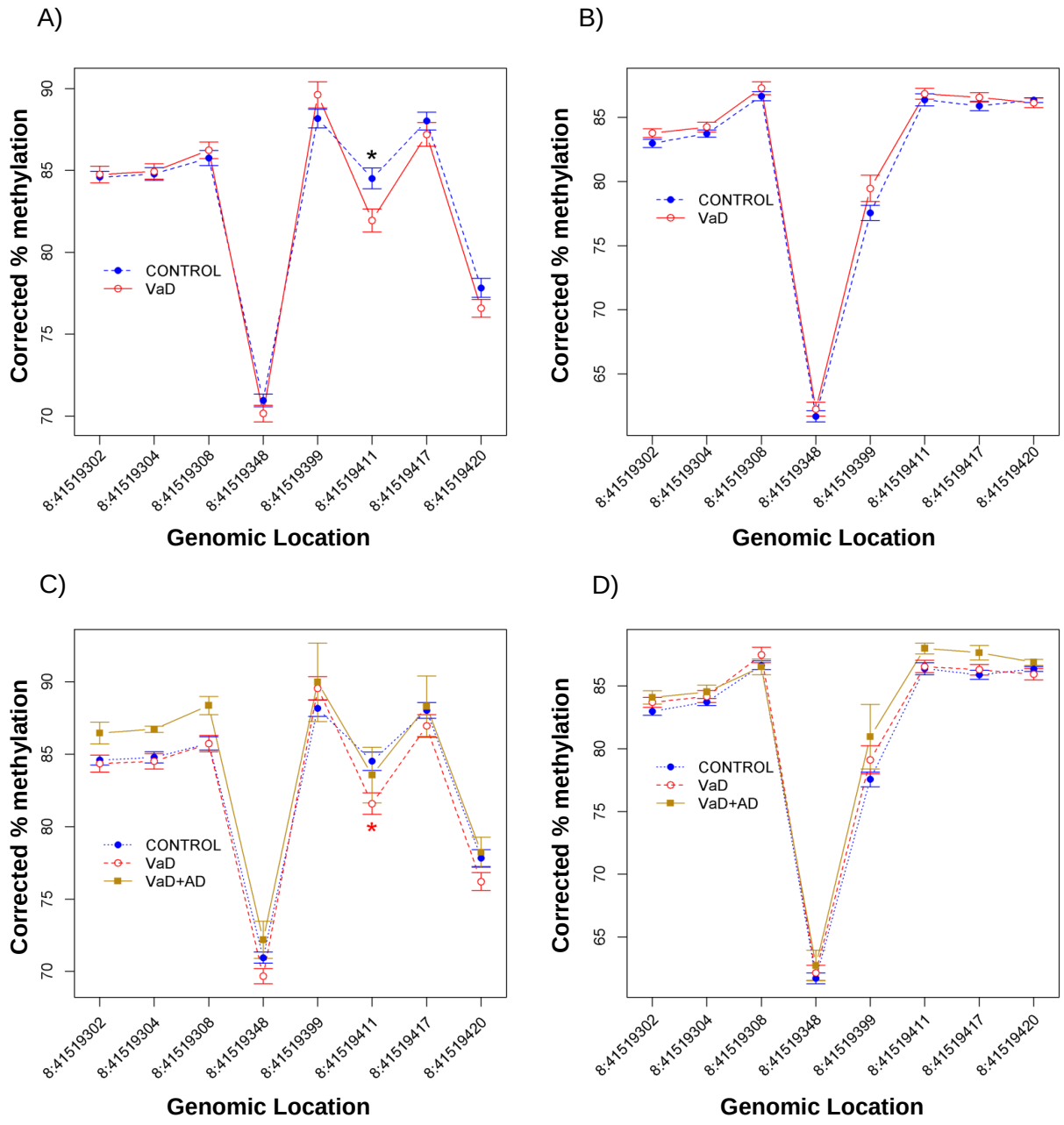


Figure 4.6; ANK1 DNA methylation levels in the STG and CER in individuals with VaD. *Using BS pyrosequencing we assayed a 118bp region of the ANK1 gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in all VaD samples compared to control samples. We observed neuropathology-associated hypomethylation at one site in the STG (A) and no disease-associated changes in the CER (B). When we compared 5mC levels in ANK1 in individuals with VaD and co-existing AD pathology, or individuals with "pure" VaD only to controls, we found significant hypomethylation at one sites in the STG in individuals with "pure" VaD (linear regression: chr8:41519411: t_{95} = -2.3, r = 0.40, P = 0.026) (C), but no disease-associated differences in the CER (D). Data is represented as mean (±SEM) Key: *P < 0.05, **P < 0.01, ***P < 0.005.*

	All VaD						VaD Only						VaD + AD					
	EC		STG		CER		EC		STG		CER		EC		STG		CER	
Genomic Location	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***
chr8:41519302	1.06	0.214	0.17	0.807	0.82	0.221	0.39	0.668	-0.24	0.742	0.72	0.322	**3.94**	**0.027**	1.87	0.180	1.11	0.429
chr8:41519304	1.38	0.127	0.16	0.827	0.52	0.404	0.66	0.492	-0.27	0.728	0.43	0.521	**4.46**	**0.020**	1.94	0.197	0.80	0.534
chr8:41519308	1.88	0.063	0.49	0.561	0.64	0.389	1.04	0.327	-0.03	0.975	0.82	0.314	**5.42**	**0.012**	2.61	0.140	0.13	0.931
chr8:41519348	0.78	0.412	-0.83	0.257	0.57	0.520	0.12	0.904	-1.29	0.092	0.44	0.647	3.62	0.063	1.23	0.406	1.03	0.577
chr8:41519399	1.47	0.278	1.49	0.176	1.96	0.132	0.52	0.716	1.36	0.233	1.56	0.261	**5.76**	**0.045**	1.79	0.437	3.40	0.206
chr8:41519411	1.53	0.068	**-2.66**	**0.026**	0.47	0.631	1.16	0.192	**-2.93**	**0.020**	0.18	0.867	3.02	0.105	-0.96	0.704	1.62	0.441
chr8:41519417	1.78	0.077	-0.84	0.418	0.69	0.351	0.89	0.394	-1.07	0.328	0.41	0.606	**5.72**	**0.008**	0.29	0.892	1.77	0.266
chr8:41519420	1.34	0.164	-1.29	0.217	-0.23	0.589	1.46	0.139	-1.62	0.146	-0.41	0.376	0.55	0.794	0.40	0.858	0.54	0.517
Amplicon Average	1.38	0.073	-0.39	0.546	0.73	0.122	0.82	0.325	-0.74	0.301	0.59	0.260	**4.06**	**0.016**	1.13	0.429	1.32	0.196

Table 4.4; ANK1 CpG sites are hypermethylated in the EC in individuals with VaD and co-existing AD pathology. *Shown for each tissue is chromosomal location (genome build hg19) of CpG sites tested, with corrected difference (Δ) in 5mC levels and corresponding P value between controls and VaD cases after adjusting for the covariates of age, sex and batch effects (linear regression). Analyses are stratified into "All VaD", "VaD Only" and "VaD + AD" groups. A positive Δ reflects hypermethylation in disease. All P values < 0.05 (and their corresponding Δ) are shown in bold.*

4.4.4 *ANK1* DNA hypermethylation in the EC is seen in both HD and PD.

ANK1 DNA hypermethylation was seen at four of the eight CpG sites in the EC in HD (Figure 4.7). However, no differential 5mC was seen in the other brain regions tested (Table 4.5), including the STG (Figure 4.8A), the CER (Figure 4.8B) and the STR (Figure 4.8C), a region that forms part of the basal ganglia, known to be the first brain region to be adversely affected by HD pathology (Walker 2007). Averaging across the region again highlighted significant hypermethylation in the EC (P = 6.68 x 10^{-3}), with no significant change in any other tissue (Figure 4.8D).

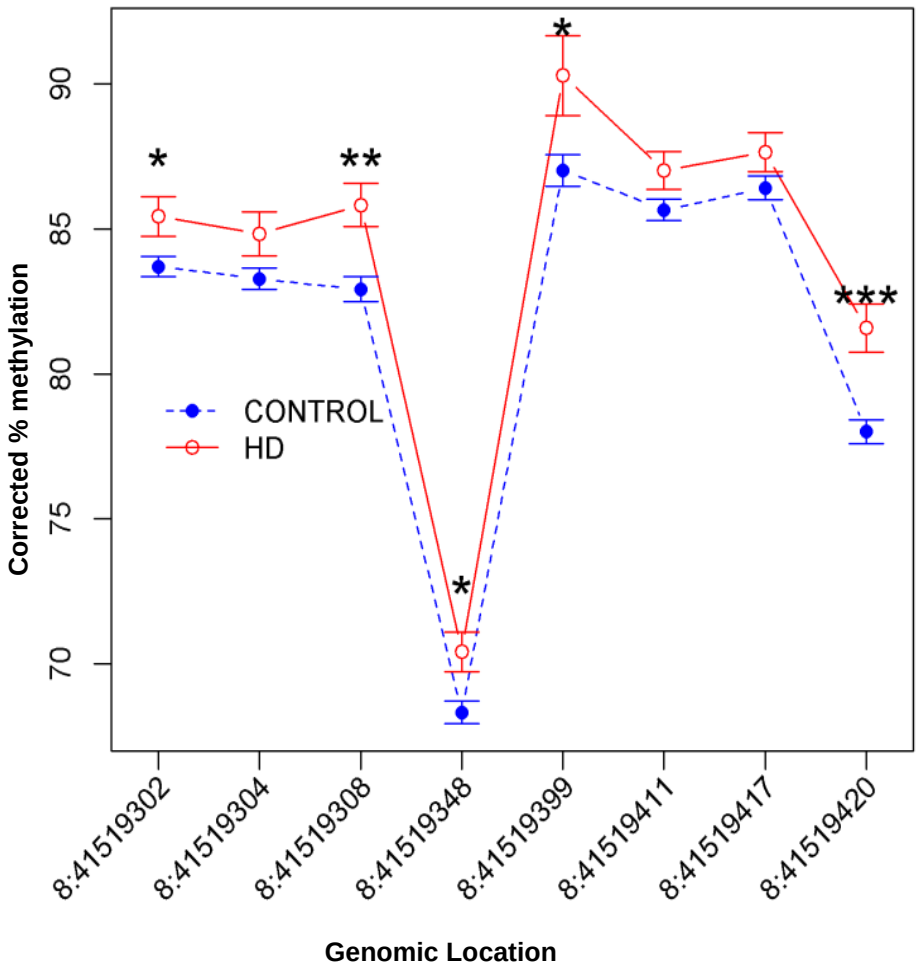


Figure 4.7; ANK1 DNA methylation patterns in HD. *Using BS pyrosequencing we assayed a 118bp region of the ANK1 gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in HD samples compared to control samples. In HD donors in the EC we observed neuropathology-associated hypermethylation at five sites (linear regression: chr8:41519302: t_{112} = 2.2, r = 0.31, P =0.027; chr8:41519308: t_{110} = 3.1, r = 0.39, P =0.003; chr8:41519348: t_{104} = 2.5, r = 0.31, P = 0.014; chr8:41519399: t_{113} = 2.5, r = 0.28, P = 0.014; chr8:41519420: t_{113} = 4.0, r = 0.40, P = $1.24x10^{-4}$). Data is represented as mean (±SEM) Key: *P < 0.05, **P < 0.01, ***P < 0.005.*

	HD							
	EC		STG		CER		STR	
Genomic Location	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***
chr8:41519302	**2.09**	**0.027**	1.20	0.160	0.33	0.681	0.76	0.496
chr8:41519304	1.89	0.062	0.58	0.532	0.31	0.666	1.14	0.377
chr8:41519308	**3.41**	**0.003**	1.61	0.136	0.99	0.261	-0.48	0.718
chr8:41519348	**2.54**	**0.014**	-0.24	0.797	-0.80	0.470	0.16	0.885
chr8:41519399	**3.89**	**0.014**	1.25	0.388	-0.68	0.673	-0.23	0.901
chr8:41519411	1.70	0.083	0.38	0.814	0.54	0.649	0.60	0.613
chr8:41519417	1.51	0.168	0.08	0.950	-0.05	0.960	0.82	0.461
chr8:41519420	**4.31**	**1.24E-04**	2.44	0.089	-0.37	0.509	1.63	0.184
Amplicon Average	**2.21**	**0.007**	0.63	0.431	0.07	0.903	0.84	0.356

Table 4.5; ANK1 CpG sites are hypermethylated in the EC in HD. *Shown for each tissue is chromosomal location (genome build hg19) of CpG sites tested, with corrected difference (Δ) in 5mC levels and corresponding P value between controls and HD cases after adjusting for the covariates of age, sex and batch effects (linear regression). A positive Δ reflects hypermethylation in disease. All P values < 0.05 (and their corresponding Δ) are shown in bold.*

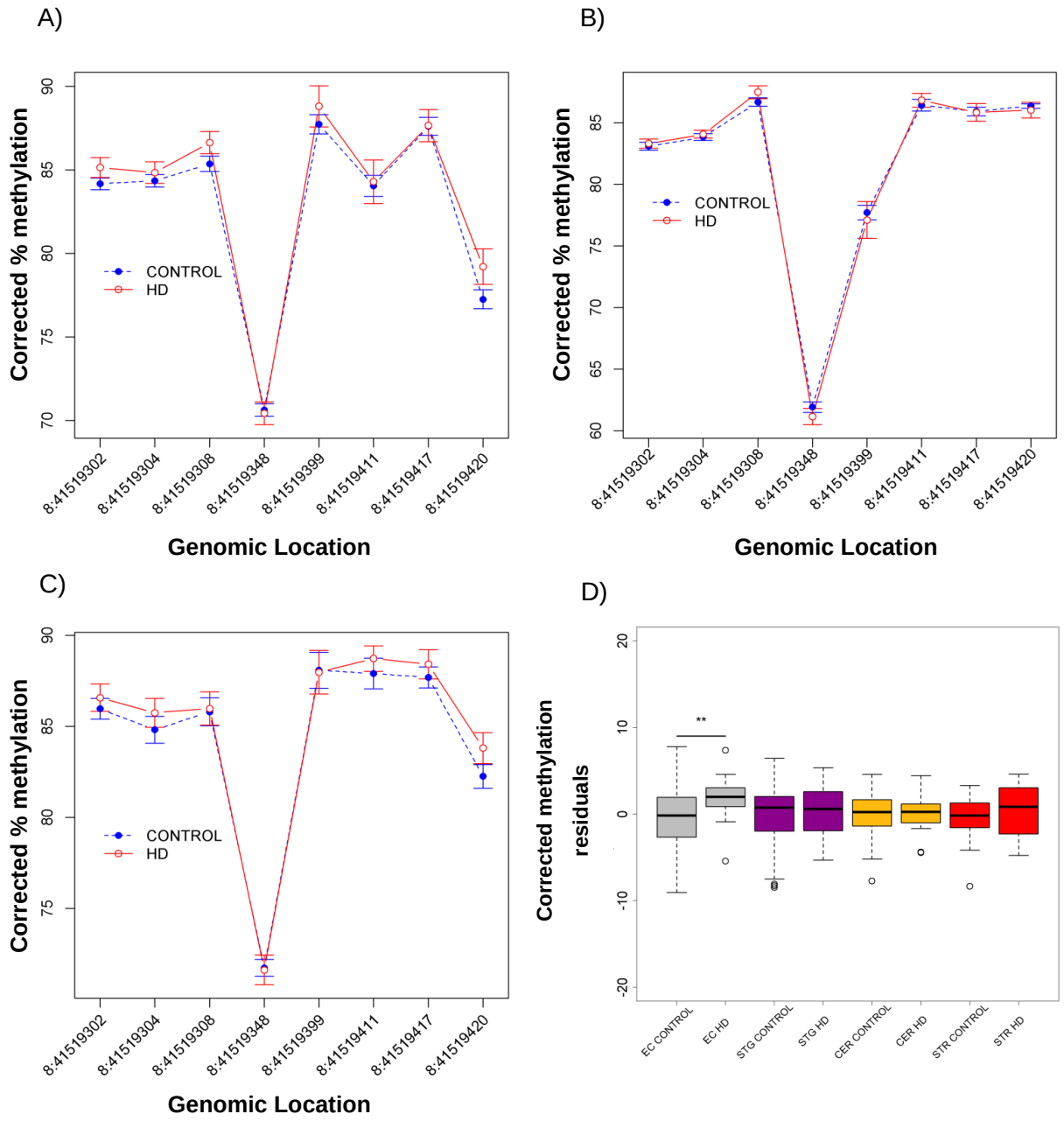


Figure 4.8; ANK1 DNA methylation levels in the STG, CER and STR in individuals with HD. *Using BS pyrosequencing we assayed a 118bp region of the ANK1 gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in HD samples compared to control samples. We did not observe any disease-associated changes in the STG (A), CER (B) or STR (C). Data is represented as mean (±SEM) Significant hypermethylation was seen in the EC when we averaged methylation across the region (two sample t-test: t_{118} = 2.8, P = 0.007), which did not meet our stringent Bonferroni significance threshold ($1.39x10^{-3}$). (D). Box and whisker plot shows median, upper and lower quartiles, and maximum and minimum values (excluding outliers). Key: *P < 0.05, **P < 0.01, ***P < 0.005.*

A similar pattern of *ANK1* hypermethylation was observed in PD; two of the eight CpG sites were characterised by significant hypermethylation in the EC (Figure 4.9A), with no differences in 5mC in any of the other brain regions tested (Table 4.6). This included the STG (Figure 4.10A), the CER (Figure 4.10B), the STR (Figure 4.10C) and the SN (Figure 4.10D), with the SN representing the brain region that has the highest levels of pathology in PD (Fearnley et al. 1991). Across the 118bp amplicon we saw no change in 5mC in any of the five brain regions (Figure 4.9B).

Genomic Location	PD									
	EC		STG		CER		STR		SN	
	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***
chr8:41519302	0.37	0.592	0.20	0.777	0.23	0.714	1.43	0.351	-0.25	0.745
chr8:41519304	0.36	0.628	0.49	0.524	-0.33	0.562	0.89	0.647	0.19	0.823
chr8:41519308	0.34	0.686	0.80	0.334	-0.55	0.518	-2.46	0.203	0.05	0.956
chr8:41519348	0.83	0.323	0.07	0.932	0.18	0.838	1.17	0.381	-0.24	0.826
chr8:41519399	0.03	0.981	-0.10	0.924	0.16	0.893	-3.40	0.200	1.37	0.272
chr8:41519411	**1.65**	**0.032**	-1.32	0.234	0.14	0.875	-0.06	0.978	0.57	0.600
chr8:41519417	0.95	0.261	-0.87	0.371	-0.42	0.551	-0.98	0.521	0.06	0.950
chr8:41519420	**1.81**	**0.026**	-1.07	0.274	0.28	0.434	-0.01	0.997	0.36	0.714
Amplicon Average	0.84	0.202	-0.28	0.670	-0.02	0.959	-0.19	0.879	0.29	0.690

Table 4.6; ANK1 CpG sites are hypermethylated in the EC in PD. *Shown for each tissue is chromosomal location (genome build hg19) of CpG sites tested, with corrected difference (Δ) in 5mC levels and corresponding P value between controls and PD cases after adjusting for the covariates of age, sex and batch effects (linear regression). A positive Δ reflects hypermethylation in disease. All P values < 0.05 (and their corresponding Δ) are shown in bold.*

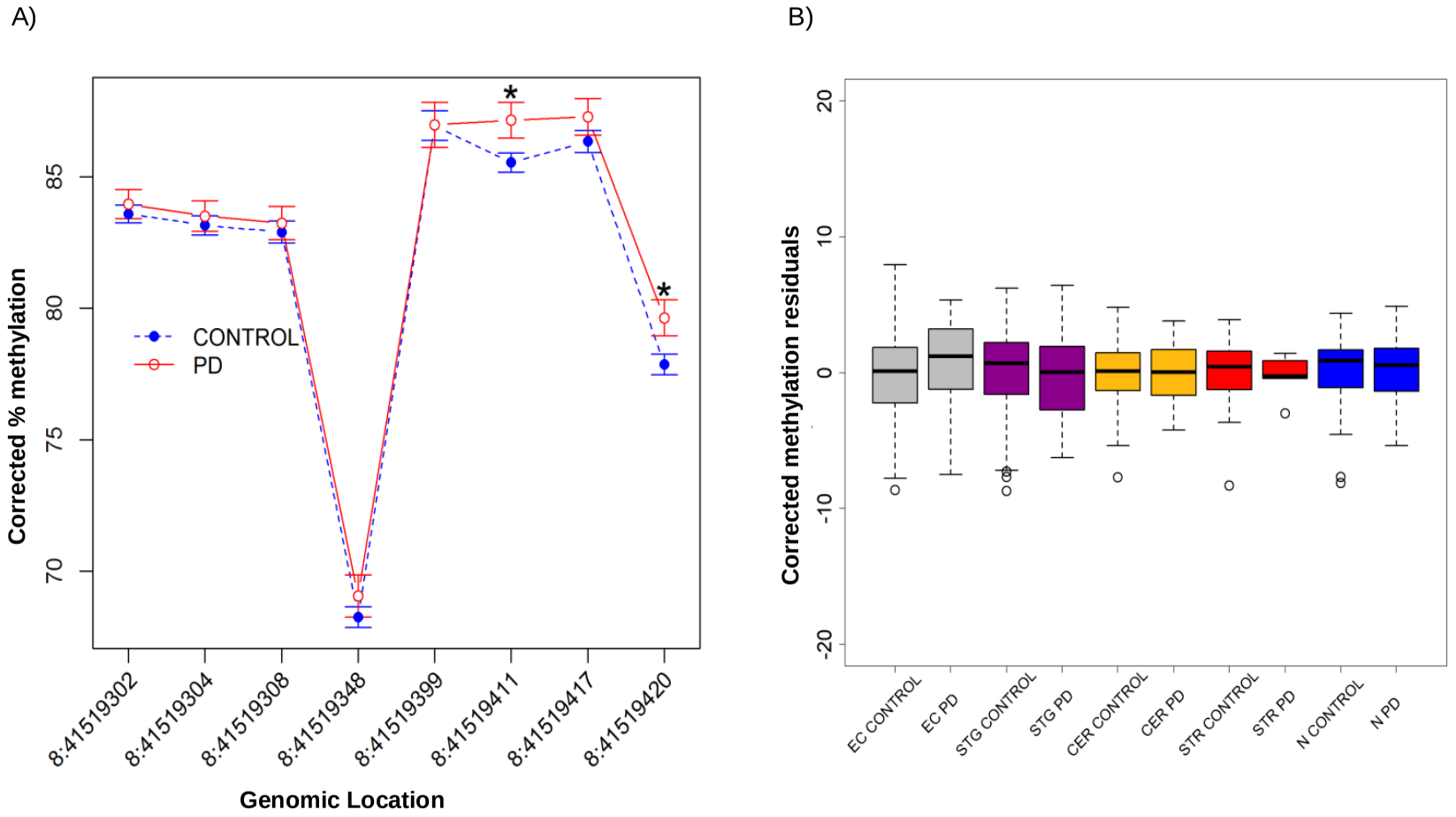


Figure 4.9; ANK1 DNA methylation patterns in PD. *Using BS pyrosequencing we assayed a 118bp region of the ANK1 gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in PD samples compared to control samples. In PD donors in the EC we found neuropathology-associated hypermethylation at two sites (A), (linear regression: chr8:41519411: t_{126} = 2.2, r = 0.30, P =0.032; chr8:41519420: t_{126} = 2.2, r = 0.35, P = 0.026), data is represented as mean (±SEM) but no significant difference when we averaged across the region (B). Box and whisker plot shows median, upper and lower quartiles, and maximum and minimum values (excluding outliers) P values were calculated using two sample t-test.Key: *P < 0.05, **P < 0.01, ***P < 0.005.*

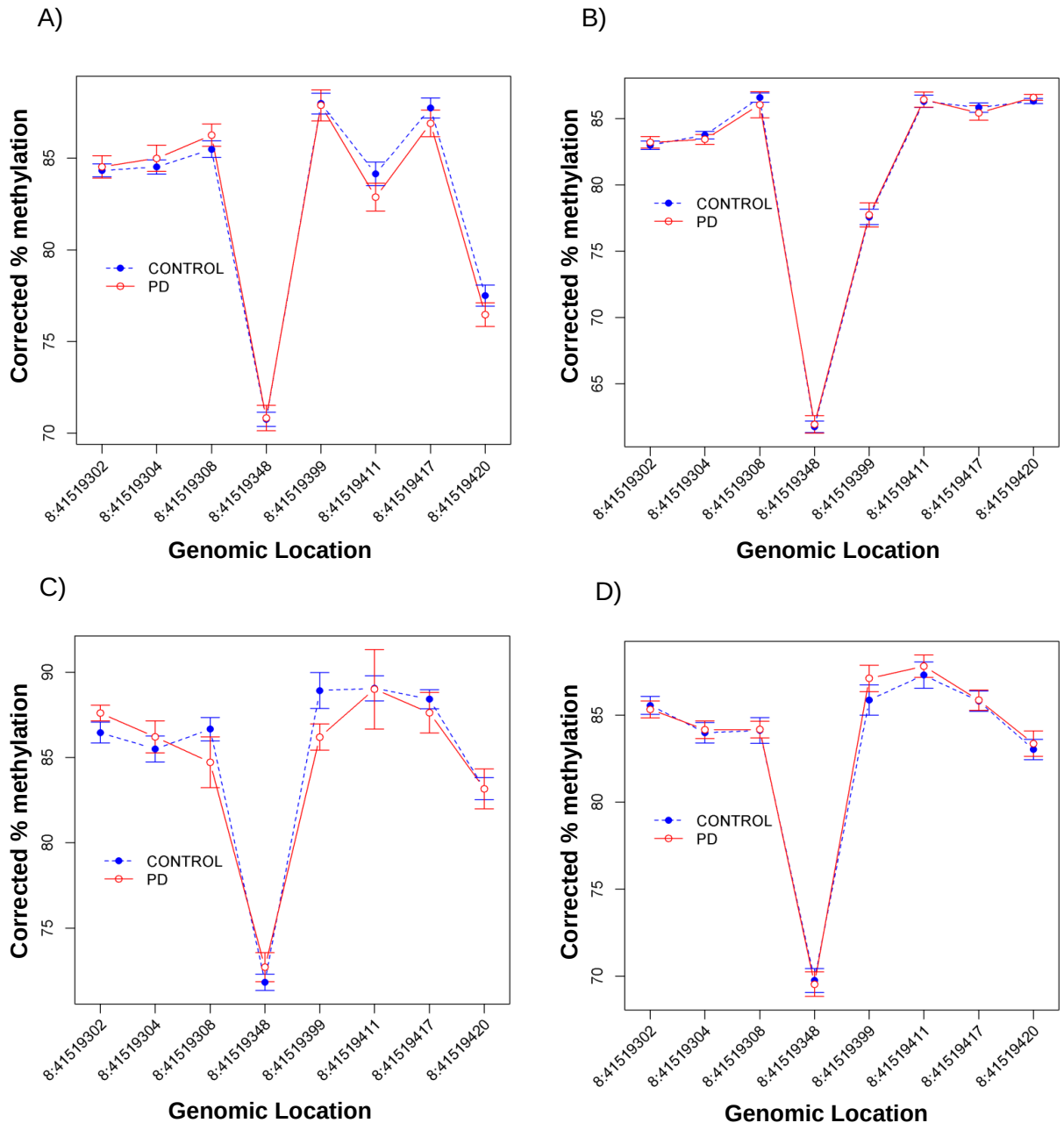


Figure 4.10; ANK1 DNA methylation levels in the STG, CER, STR and SN in individuals with PD *Using BS pyrosequencing we assayed a 118bp region of the ANK1 gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in PD samples compared to control samples. We did not observe any disease-associated changes in the STG (A), CER (B), STR (C) or SN (D), P values were calculated by linear regression. Data is represented as mean (±SEM) Key: *P < 0.05, **P < 0.01, ***P < 0.005.*

4.5 Discussion

This is the first study to assess brain *ANK1* 5mC differences across multiple neurodegenerative diseases. We identified significant 5mC differences in the EC in multiple diseases, including AD, HD and PD, with significant DNA hypermethylation across the amplicon in AD and HD. Interestingly, we also observed significant hypermethylation of *ANK1* in the EC in both DLB and VaD at several individual CpG sites and across the amplicon, but only in donors with co-existing AD pathology. This suggests that *ANK1* DNA hypermethylation in the EC is specific to some neurodegenerative diseases (AD and HD), and not observed in other forms of neuropathology (VaD and DLB). We have previously reported that *ANK1* is not differentially methylated or hydroxymethylated in the CER in AD at two sites interrogated by the Illumina 450K array (chr8:41519308 and chr8:41519399) (Lunnon et al. 2014) (Chapter 3). In the current study we again demonstrate that these two loci are not significantly differentially methylated in AD CER; however, we do highlight AD-associated DNA hypermethylation at five adjacent CpG sites and averaged across the 118bp amplicon in the CER. We did not see any *ANK1* 5mC changes in the CER in any of the other neurodegenerative diseases, including those with co-existing AD pathology. Reflecting our previous findings, we found that 5mC differences in AD are greatest in the EC, an area with high levels of neuropathology and lowest in the CER, the region with the least neuropathology. Interestingly, although we observed disease-associated *ANK1* hypermethylation in the EC at six individual sites and across the region in HD and at two individual sites in PD, we did not see *ANK1* 5mC changes in these diseases in their regions of primary neuropathology, namely the STR and SN, respectively. This suggests that *ANK1* hypermethylation in neurodegenerative disease is not necessarily specific to regions of primary neuropathology, but may instead be specific to particular cell types affected in the diseases, such as those in the EC, which are not present in the STR and SN. One caveat of our study is that we have analysed “bulk” tissue, and we cannot determine which cell type(s) are driving the DNA hypermethylation we see in *ANK1* in disease. *ANK1* encodes for numerous isoforms with their own tissue-specific enhancers. Although the precise function of most *ANK1* isoforms is not known, different isoforms have been identified in the brain, blood and muscle (Gallagher et al. 1997). Mastroeni *et al*., recently showed a four-fold increase in

ANK1 mRNA expression in microglia from AD brain tissue, but not in neurones or astrocytes from the same individuals, suggesting an immune based function for *ANK1* in the human brain (Mastroeni et al. 2017).

Although the focus of our study was on investigating 5mC changes in neurodegenerative disease, BS pyrosequencing actually generates a summative measurement of both 5mC and 5hmC. 5mC is generally associated with gene silencing, particularly when residing in the promoters of constitutively expressed housekeeping genes (Jones 2012), whilst 5hmC has been shown to be enriched in gene bodies (Lunnon et al. 2016) and to be found at (relatively) high levels in the brain (Szulwach et al. 2011, Khare et al. 2012). In Chapter 3 we have shown that *ANK1* DNA hypermethylation across the 118bp amplicon in AD is actually underestimated when using BS data, as it is confounded by significant DNA hypohydroxymethylation at some loci in the amplicon (Smith et al. 2018). Another potential caveat of our study ; is that we have only analysed 5mC across eight CpG sites in an 118bp region of the *ANK1* gene and thus future studies should aim to further quantify changes in 5mC across the entire 244kb gene.

4.6 Conclusions

Our study has demonstrated disease-associated *ANK1* hypermethylation in the EC at specific CpG sites in AD, HD and PD and across the region in AD and HD. In donors with DLB and VaD we only observed increased *ANK1* 5mC in the EC only in individuals with co-existing AD pathology. The CER showed disease-associated hypermethylation at specific CpG loci and across the region in AD, but not in any of the other neurodegenerative diseases tested. We saw no disease-associated differential *ANK1* 5mC in the STR in HD or PD, or the SN in PD. This suggests that *ANK1* is characterised by brain region and disease-specific differential 5mC in different neurodegenerative diseases.

CHAPTER 5: EXPLORING HISTONE MODIFICATIONS AT THE *ANK1* LOCUS

The work presented in this chapter has not yet been published. However, we are in the process of writing up this data for publication.

5.1 Introduction

Histone proteins were first discovered by Albrecht Kossel in 1884, but were initially dismissed as transcriptional regulators as they were thought to be only an inert packaging material. It was Kayne et al. (1988) who demonstrated that histones innately repress transcription *in vivo*. This gave rise to the theory that modifiers of histone structure could have an effect on gene transcription. Today it is known that histone modifications represent the reversible short-term regulation of hetero/euchromatin (Cedar et al. 2009). A large number of histone modifications have been described, with many predicted functions, for example the di-methylation of the ninth lysine on histone 3 (H3K9me2) is a marker of gene repression (Kim 2014). Two of the most studied histone modifications are the tri-methylation of lysine at positions 4 and 27 of histone 3 (H3K4me3 and H3K27me3), which are markers of gene expression and repression respectively. These modifications are the focus of this chapter.

5.1.1 H3K4me3

H3K4me3 is a chromatin modification that is consistently present at transcription start sites (TSS) of actively transcribed genes in eukaryotes. High levels of the modification reflect the amount of transcription, which was initially described by Santos-Rosa et al. (2002). Together with other reports (Azuara et al. 2006, Kouzarides 2007, Mattout et al. 2010), this has led to the hypothesis that H3K4me3 is a transcription 'activating' histone modification. Mono and di-methylated forms of this modification are also thought to be markers of actively transcribed regions, although there is less research into these modifications. H3K4me3 is also localised at the 5' end of active genes near the TSS and was found to be associated with the initiated form of RNA polymerase II (Barski et al. 2007). However, a report has shown widespread decreased H3K4me3 at multiple expressed genes with only minimal changes in the transcript abundance of these genes (Clouaire et al. 2012).

This suggests that H3K4me3 may not be the sole driver of gene transcription, with other epigenetic factors affecting transcription. Alternatively this may be in part due to the cell-specific nature of epigenetic changes, meaning that such changes may only be detectable at single cell resolution. It is also possible that this modification is deposited as a result of transcription rather than driving it.

5.1.2 H3K27me3

Unlike H3K4me3, H3K27me3 is associated with inactive gene promotors, causing a dramatic and predictable lack of gene expression (Boyer et al. 2006, Lee et al. 2006, Barski et al. 2007). It is involved in the maintenance of the inactive X-chromosome silencing as well as being involved in genomic imprinting (Kouzarides 2007). H3K27 can undergo a number of other modifications (Table 1.2). Lysine at position 27 of histone 3 can also exist in mono- or di-methylated states, although the role of these other modifications are not as well characterised as H3K27me3. In addition, it is known that H3K27me3 is propagated during cellular division; it is suggested that H3K27me3 recruits a Polycomb Repressive Complex 2 (PRC2) that in-turn leads to the methylation of H3 histones on the newly synthesised DNA (Hansen et al. 2008). H3K4me3 and H3K27me3 can occur on the same histone tail: these so-called “bivalent modifying patterns” or “primed chromatin” mean that viewing modifications as “activating” or “silencing” is overly simplistic(Lesch et al. 2014). These bivalent patterns in embryonic stem cells (ESC) are predicted to cause highly dynamic levels of gene expression (Bernstein et al. 2005).

5.1.3 Histone modifications and their relationship to DNA modifications

5mC is generally associated with gene silencing when it occurs in the promoter regions of genes (Curradi et al. 2002). However, at other genomic locations, for example gene bodies, 5mC is believed to result in gene activation or alternative splicing (Yang et al. 2014). There is evidence of cross-talk between DNA and histone modifications (Kondo 2009). For example, during embryo implantation when the epigenetic landscape within each cell is being re-programmed, methylation at CpG islands is protected from removal, through a mechanism that is dependent on the

recognition of H3K4 methylation. In addition 5mC acts as a template for histone modification patterns after DNA replication, mediated through methylcytosine-binding proteins that are capable of recruiting histone deacetylases (Nan et al. 1998, Chi et al. 2010) and methyltransferases (Jin et al. 2011). This crosstalk can be caused by biochemical interactions between the SET domain (an evolutionary conserved sequence motif) of histone methyltransferases which are responsible for the methylation of histones and the recruitment of DNA methyltransferases (Cedar et al. 2009).

5.1.4 Histone modifications in neurological disease

Similarly to DNA modifications, histone modifications have been associated with neurological diseases, such as AD. For example in amyotrophic lateral sclerosis (ALS) the acetyltransferase *CREB* binding protein (CBP) is inhibited by the fused in sarcoma (*FUS*) protein. This in turn leads to histone hypo-acetylation. Similar mechanisms are seen in PD, HD and Friedreich's ataxia (Urdinguio et al. 2009). Interestingly, AD has also been linked to incorrect modification of histone proteins. Cleavage of the APP protein in AD leads to the production of AICD (Figure 1.1) which, coupled with Fe65, activates TIP60, another histone acetyltransferase. This leads to the increased acetylation of H3K14 and H4K5 (Kim et al. 2004). In addition, work in mice has revealed that neuron-specific overexpression of histone deacetylase 2 (HDAC2) decreases dendritic spine density, synapse number, synaptic plasticity and memory function. Whereas HDAC2 deficiency causes an increase in synaptic number and memory facilitation (Guan et al. 2009). Recent work has profiled acetylation of the twenty seventh lysine on histone 3 (H3K27ac) across the genome in post-mortem tissue, identifying widespread AD-associated variation in this histone modification. Differentially acetylated peaks were identified in genomic regions associated with both tau and amyloid neuropathology and variants associated with late-onset AD (Marzi et al. 2017). This further supports a role for alterations in histone modifications in neurological diseases, such as AD.

To date only one study has examined histone modifications in *ANK1*, by interrogating the ENCODE Chromatin Immunoprecipitation Sequencing (ChIP-seq) dataset (Bernstein et al. 2005). This study focused on the proposed erythrocytic

specific promotor of *ANK1* and the level of enhancer associated histone modifications in this region (H3K4me1 and H3K27ac) (Yocum et al. 2012). However, this approach was limited to K562 cells (human blood cell line), hence the applicability of this data to brain tissue is limited.

5.1.5 Chromatin Immunoprecipitation

ChIP has been widely used to map the location of modified histones, transcription variants, transcription factors and many other DNA-protein interactions. ChIP is the use of an antibody to target and isolate a protein or histone modification of interest and then enrich regions of DNA with which it associates. In recent years this technique coupled with micro-array (ChIP-chip) and high-throughput sequencing technologies (ChIP-seq) has enabled the profiling of histone modifications on a genome-wide scale (Collas 2010). Alternatively, the coupling of ChIP with qPCR (ChIP-qPCR) allows the characterisation of histone modifications at a specific candidate locus.

5.1.6 Summary

Although we have now shown that 5mC and 5hmC changes are seen in the *ANK1* gene AD brain, in Chapters 3 and 4. At present no study has yet examined alterations in histone modifications at the *ANK1* locus in AD brain. This chapter aims to utilise ChIP-qPCR to quantify H3K4me3 and H3K27me3 modification changes in the *ANK1* gene in post-mortem brain samples from donors with varying degrees of AD pathology.

5.2 Aims

- To optimise and use ChIP-qPCR to measure H3K4me3 and H3K27me3 in *ANK1* in human, post-mortem brain samples.
- To use the optimised ChIP-qPCR assays on matched brain samples to those used for DNA (hyroxy)methylomic profiling in Chapter 3.
- To investigate whether *ANK1* H3K4me3 and H3K27me3 levels are altered in the EC of individuals with high levels of neuropathology compared to those with low levels.
- To examine whether *ANK1* H3K4me3 and H3K27me3 levels are correlated with each other and DNA modification levels in post-mortem brain samples.
- To investigate whether *ANK1* H3K4me3 and H3K27me3 levels are correlated with 5mC and 5hmC levels at key loci previously shown to be differentially methylated in AD EC (cg11823178 and cg05066959)

5.3 Materials and Methods

5.3.1 Subjects and samples

This study used EC tissue from post-mortem brain samples collected from 88 individuals archived in the LNDBB. 63 of these samples were utilised for DNA methylomic and hydroxymethylomic profiling in Chapter 3. Donors used for the study had varying degrees of AD pathology, ranging from Braak Stage 0 to Braak Stage VI, and were all over the age of 65. All samples were dissected by trained specialists, snap-frozen and stored at -80°C. Further demographic information about all samples is provided in Table 5.1.

	Low pathology	High pathology
Number of samples	29	59
Sex (%) M/F (%)	(48) 14/15 (52)	(59) 35/24 (41)
Mean Age (±SD)	79.2 (11.2)	82.2 (8.2)
Braak Stage	0-III	IV-VI

Table 5.1; Demographic information for EC brain samples used in this chapter. *The distribution of sex, Braak stage and mean age (± standard deviation (SD)) are provided.*

5.3.2 ChIP for H3K4me3 and H3K27me3

H3K4me3 and H3K27me3 profiling was performed using methodology documented in Chapter 2, Section 2.10.1. Briefly, frozen post-mortem brain tissue samples were cut into 10-15mg sections as described in Section 2.2. Following this, DNA and associated proteins were cross-linked using formaldehyde and cells were then re-suspended in HBSS (Section 2.10.1.2). Cells were incubated in lysis buffer (Diagenode, cat no.: C01010130) and sonicated for 35 minutes to shear the chromatin (Section 2.10.1.3). After fragmentation each chromatin sample was analysed via gel electrophoresis (Section 2.7.2) to check the size of the chromatin fragments were between 100-600bp (Figure 5.1). Following this, automated magnetic immunoprecipitation was used for each chromatin sample, using three separate antibodies (100µL of chromatin per antibody): H3K4me3, H3K27me3 and an IgG negative control, as described in Section 2.10.1.4. Finally, DNA was de-crosslinked from the protein and was purified using spin columns (Diagenode, cat no.: C01010130) (Section 2.10.1.5) and eluted in 80µL leaving pure DNA for downstream qPCR.

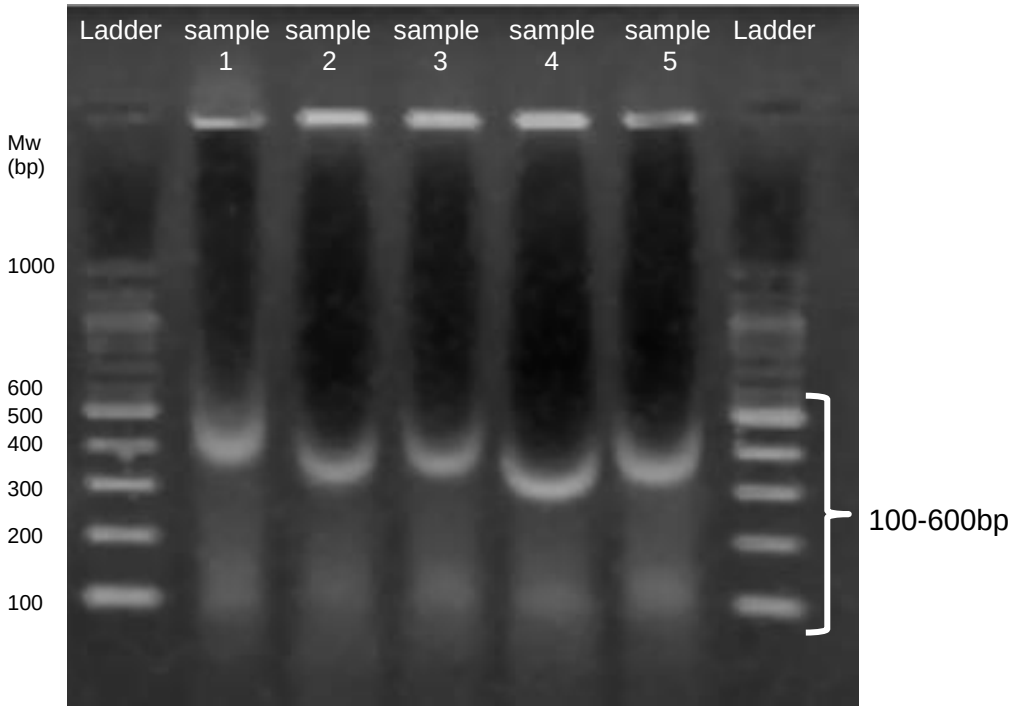


Figure 5.1; Chromatin fragmentation gel. *An example gel image of fragmented chromatin for samples used in the study. Five of the samples used in the study are shown alongside a 100bp ladder (Solis BioDyne cat no.: 07-11-00050). All samples were fragmented to between 100-600bp before qPCR.*

5.3.3 ChIP-qPCR

qPCR was used to detect specific DNA target sequences in *ANK1.* 2µL ChIP treated DNA was combined with 2µL EvaGreen master mix (SolisBiodyne, cat no.: 08-24-00008), 2-3 µM forward and reverse primer sequences (Table 5.2) and 5µL of H_2O per sample replicate in 384 well plates (Section 2.11.1, Table 2.20). Seven 1 in 4 serial dilutions and a NTC were included in triplicate per plate. The resulting plate was run on a QuantStudio (Thermo Fisher, US) qPCR machine with the following settings: 95°C for 10 minutes, 45 cycles of: 95°C for 30 seconds, 60°C for 30 seconds, 72°C for 30 seconds, followed by a dissociation stage.

5.3.4 *ANK1* ChIP-qPCR primer design

ChIP-qPCR primers were designed using Primer3Plus software (Untergasser et al. 2007). Primers were designed (a) to target loci identified in Chapter 3 and 4 as differentially methylated in AD (chr8: 41519302 to 41519420), (b) across the remainder of the *ANK1* gene in known areas of high H3K4me3 and H3K27me3 abundance and (c) to other regulatory regions of the *ANK1* gene (see Figure 5.2). In total 14 ChIP-qPCR assays were designed and used (Table 5.2). Following quality checks and optimisation, it was decided that seven primer sets were not suitable for inclusion in the final study. Exclusion criteria included: failure to produce reproducible Ct. Scores (within 0.5), primer efficiency of less than 85% or greater than 115% and the detection of secondary products via melt curve analysis and gel electrophoresis. In total six *ANK1* and one control primer set (MyoD1) (Taberlay et al. 2011) were used in the study.

Primer Name	Genome Coordinates	Position	qPCR Primers (5' to 3')	TM (°C)	Product length (bp)	Efficiency	Included	Reason for exclusion	Final Concentration
ANK1 Primer Set 1	chr8:41510971-41511059	F	CCAAACTCCTCCCCTAACG	59.5	89	54%	-	low efficiency	-
		R	AGCTCCAGACTCCGCTTTG	60.7					
ANK1 Primer Set 2	chr8:41516703-41516777	F	TAAGGAACATCCCAGCCTTC	59.1	75	87%	✓	-	2 µM
		R	TGTTTAGAATGGCCCCTGAG	60.1					
ANK1 Primer Set 3	chr8:41559556-41559664	F	AGAGGCAGACATACCACGAAAC	60.6	109	77%	-	low efficiency	-
		R	AGGATGGAACCACACCTCTG	60					
ANK1 Primer Set 4	chr8:41570640-41570726	F	AGGTGTTGGCTGGAGTCAC	58.7	87	-	-	inconsistent Ct. scores	-
		R	TATGGGAGATCCGTCATGTG	59.3					
ANK1 Primer Set 5	chr8:41625416-41625491	F	CCGATTTGCATGTTTTGAGC	61.5	76	93%	✓	-	2 µM
		R	CTCAGGCAGAATCTGGAATAGG	60.2					
ANK1 Primer Set 6	chr8:41649940-41650067	F	TCCTAACAGCCAAAGCTCAC	58.1	128	72%	-	low efficiency	-
		R	TTCGACCAAGGTCTTGAACC	60.1					
ANK1 Primer Set 7	chr8:41686181-41686328	F	TGGTTCGCACTACAATGACC	59.6	148	97%	✓	-	2 µM
		R	TCGCGTGGAAAATCTTCAG	59.9					
ANK1 Primer Set 8	chr8:41753871-41753986	F	ACCTTCTTCTTTCGGTCACG	59.3	116	80%	-	low efficiency	-
		R	GCCAAACAGCTGAAGAAAATC	59					
ANK1 Primer Set 9	chr8:41754877-41755012	F	GAGCTTGATTTGCCCCTTTC	61.1	136	100%	✓	-	2 µM
		R	CGGGGACCTACTTCCTTTTC	59.9					
ANK1 Primer Set 10	chr8:41519214-41519359	F	GAGCTCATAAAGAGGTGAGTGG	58.1	146	87%	-	Lower efficiency than primer 11, covering same genomic region	-
		R	AGCTGGAGGTCGATATTGATTAC	58.7					
ANK1 Primer Set 11	chr8:41519216-41519359	F	GCTCATAAAGAGGTGAGTGGAG	58.1	144	89%	✓	-	2 µM
		R	AGCTGGAGGTCGATATTGATTAC	58.7					
ANK1 Primer Set 12	chr8:41519342-41519460	F	CAATATCGACCTCCAGCTCAC	59.7	119	95%	✓	-	3 µM
		R	GATCATTCGCAAGGTGGTTC	60.5					
ANK1 Primer Set 13	chr8:41519342-41519483	F	CAATATCGACCTCCAGCTCAC	59.7	142	98%	-	secondary product	-
		R	TTCCTCCTTCCCTTCCTTTC	59.6					
*MyoD*1 Primer Set 1	chr11:17741321-17741428	F	CGCCAGGATATGGAGCTACT	59.3	108	95%	✓	-	2 µM
		R	CGGGTCGTCATAGAAGTCGT	60.1					

Table 5.2; ChIP-qPCR primer design. *Thirteen ChIP-qPCR primer sets were designed across the ANK1 gene, covering known H3K4me3 and K3K27me3 peaks. ChIP-qPCR primers to MyoD1 were used as a control set. The MyoD1 primer design was obtained from (Taberlay et al. 2011).*

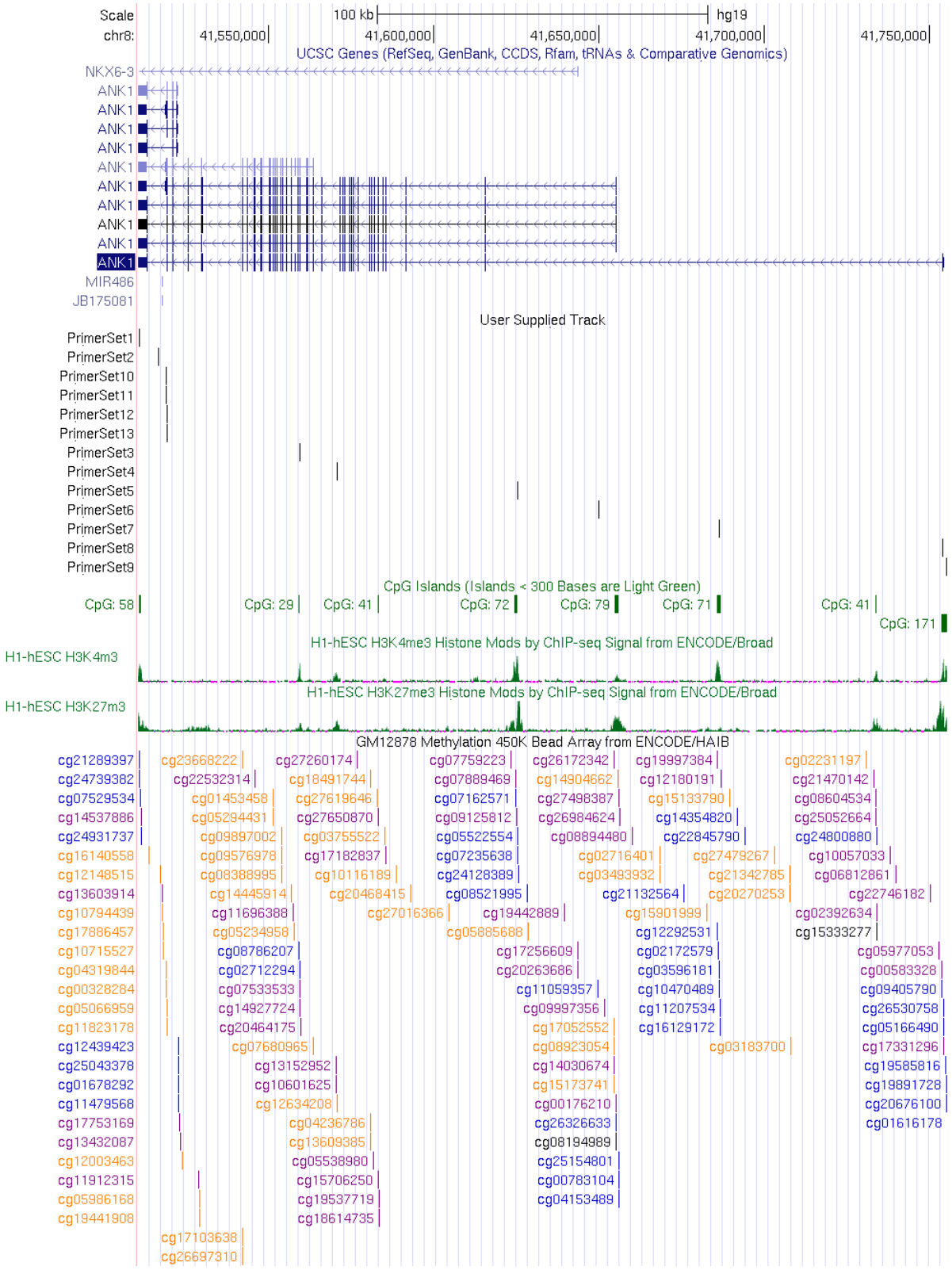


Figure 5.2; Visualisation of ChIP-qPCR primer sets designed for this study. *In total, thirteen ChIP-qPCR primer sets were designed across the ANK1 gene, covering known H3K4me3 and K3K27me3 peaks. Displayed is the complete ANK1 gene with all 10 transcript variants (shown in purple), locations of all 13 primer sets (shown in black) established H3K4me3 and H3K27me3 profiles in ESCs (shown in green), as well as the location of CpG sites covered by the 450K array including cg05066959 & cg11823178 (shown in orange).*

5.3.5 Data analysis

All samples were run in triplicate for each assay and averaged to collect a mean Ct score for each modification. A range of 0.5 Ct maximum was allowed between triplicate samples; Ct scores that fell outside this range were excluded and then an average of the two remaining scores was used. A control assay (MyoD1) with a known histone modification profile was also run for all samples (Figure 5.3). In addition a negative control IgG ChIP reaction was run for all samples in all assays, as per the manufacturer's recommendations. Standard curves of a 1 in 4 dilution series were run on each plate and samples were normalised according to primer efficiency as described in Section 2.8.2.

All computations and statistical analyses were performed using R 3.3.2 (Vienna 2012). A linear regression analysis was performed, accounting for the effects of age and sex, comparing samples with a low neuropathological burden (Braak 0-III) with samples with a high neuropathological burden (Braak IV-VI) for each primer region. We then used a two sample t-test to assess the difference in the mean of examined histone modification levels averaged across all primer regions between low (Braak 0-III) and high (Braak IV-VI) neuropathological burden. Bonferroni threshold of $P < 0.017$ was used to correct for multiple testing. A subset (N = 63) of samples used in this chapter had matched DNA methylomic and hydroxmethylomic data as collected in Chapter 3. A Pearson's correlation was used to compare histone modification levels with 5mC and 5hmC levels in *ANK1* as derived from the Illumina 450K array data generated in Chapter 3.

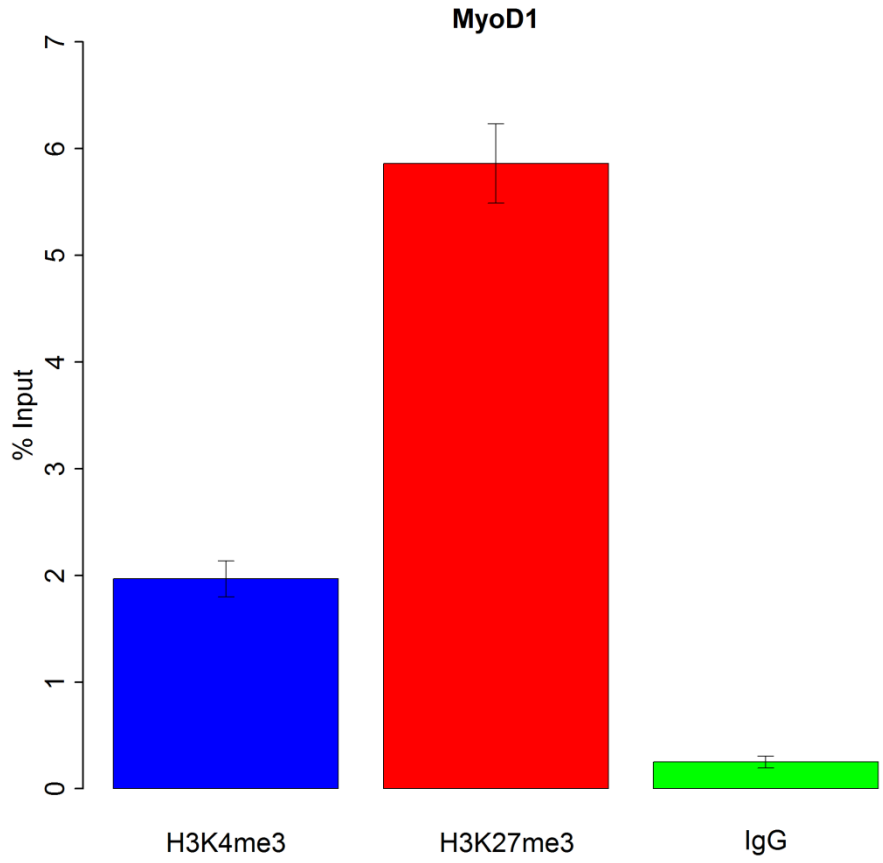


Figure 5.3; MyoD1 shows high H3K27me3 levels. *MyoD1 was used as a positive control for H3K27me3 as previous studies have shown this to be a genomic region with high modification peaks. As expected we found high levels of H3K27me3 in this region. Data is represented as a percentage relative to input*

5.4 Results

5.4.1 *ANK1* has decreased levels of H3K4me3 in individuals with high neuropathology

We assessed levels of H3K4me3 and H3K27me3 across the *ANK1* gene in AD EC using the six qPCR assays that passed QC. Of the six regions assayed, four showed a significant decrease in the level of H3K4me3 between individuals with low and high AD neuropathological burden (Primer Set 5 (chr8:41625416-41625491): Figure 5.4B, Δ = -1.60, *P* = 0.004; Primer Set 7 (chr8:41686181-41686328): Figure 5.4C, Δ = -2.53, *P* = 0.023; Primer Set 9 (chr8:41754877-41755012): Figure 5.4D, Δ = -0.41, *P* = 0.002; Primer Set 12 (chr8:41519342-41519460), Figure 5.4F, *P* = 0.033), with no significant change in the other two regions tested (Figure 5.4A, Figure 5.4E). We observed no significant difference in either H3K27me3 or the IgG control between individuals with low and high levels of pathology, in any of the six regions examined (Figure 5.4; Figure 5.5; Table 5.3). Averaging the histone modification profiles across all six regions tested demonstrated significantly decreased H3K4me3 in cases with high AD neuropathology compared to those with low neuropathological burden (Δ = -0.98, *P* = 0.004), with no change in H3K27me3 or the IgG control (Figure 5.6).

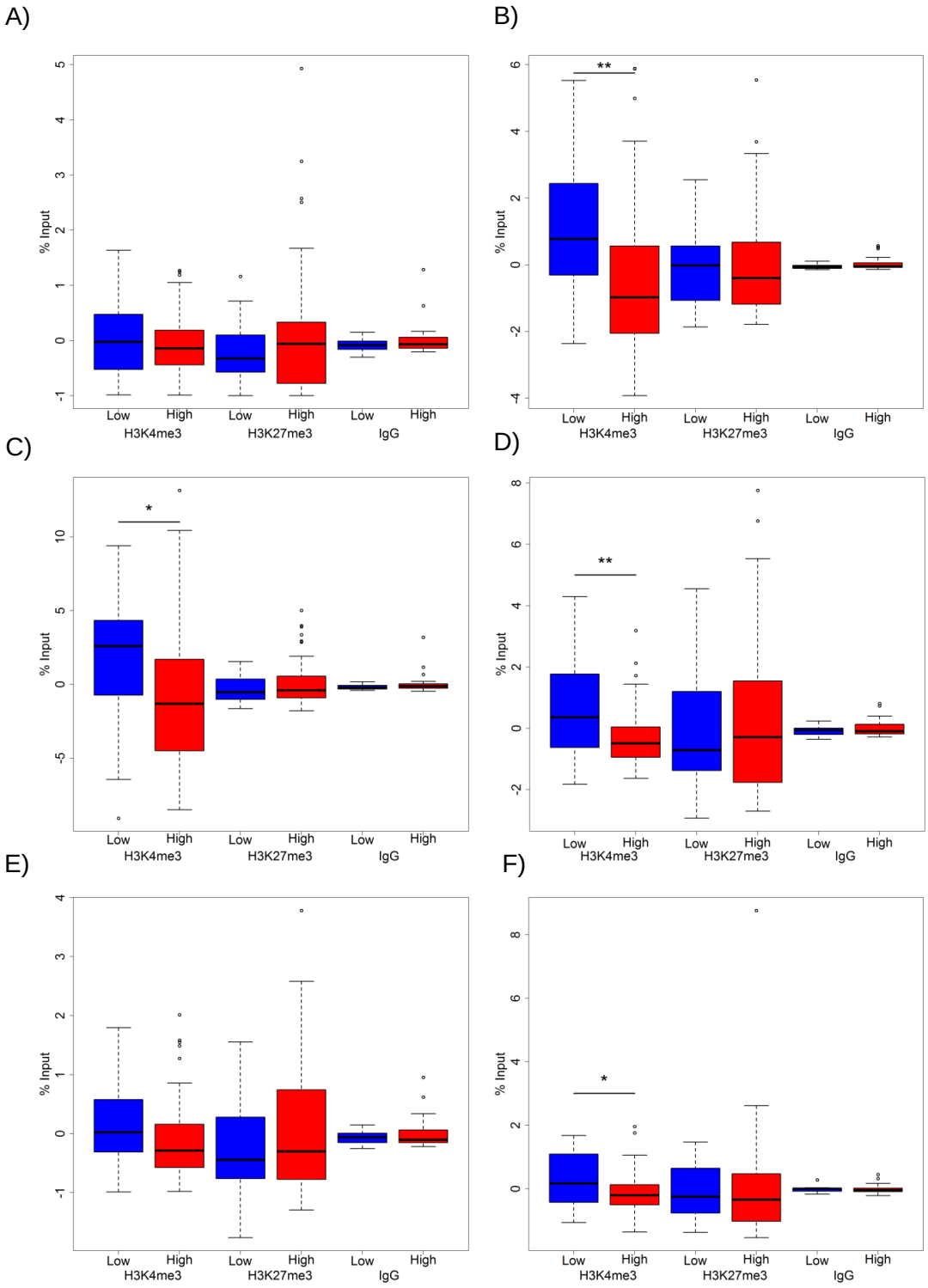


Figure 5.4; H3K4me3 levels are decreased in AD EC in specific regions of the ANK1 gene. *We observed decreased H3K4me3 in individuals with high levels of pathology compared to those with low levels of pathology at (B) chr8:* 41625416-41625491*: (primer set 5) (t_{82} = -2.9, r = 0.29, P = 0.004), (C) chr8: 41686181-41686328 (Primer set 7) (t_{83} = -2.3, r = 0.23, P = 0.023) (D) chr8: 41754877-41755012 (Primer set 9) (t_{80} = -3.2, r = 0.32, P =0.002) and (F) chr8: 41519342-41519460 (Primer set 12) (t_{67} = -2.2, r = 0.37, P =0.033) (linear regression). We found no significant difference between AD cases and controls for H3K27me3 or the IgG control for any of the brain regions tested. Data is represented as a percentage relative to input, box and whisker plots show median, upper and lower quartiles, and maximum and minimum values (excluding outliers). Key: *P < 0.05, **P < 0.01, ***P < 0.005.*

Genomic region	Modification	Δ	*P*
chr8:41516703-41516777 (Primer set 2)	H3K4me3	-0.09	0.541
	H3K27me3	0.36	0.138
	IgG	0.15	0.228
chr8:41625416-41625491 (Primer set 5)	H3K4me3	**-1.60**	**0.004**
	H3K27me3	0.02	0.949
	IgG	0.10	0.096
chr8:41686181-41686328 (Primer set 7)	H3K4me3	**-2.53**	**0.023**
	H3K27me3	0.49	0.178
	IgG	0.33	0.243
chr8:41754877-41755012 (Primer set 9)	H3K4me3	**-0.91**	**0.002**
	H3K27me3	0.08	0.884
	IgG	0.13	0.383
chr8:41519216-41519359 (Primer set 11)	H3K4me3	-0.25	0.160
	H3K27me3	0.22	0.394
	IgG	0.11	0.361
chr8:41519342-41519460 (Primer set 12)	H3K4me3	**-0.41**	**0.033**
	H3K27me3	0.11	0.770
	IgG	0.01	0.947

Table 5.3; H3K4me3 levels are decreased in specific regions of the ANK1 gene in AD EC. *Shown for each genomic region assayed is the corrected difference (Δ) in histone modification levels relative to input and corresponding linear regression derived P value between individuals with low levels of pathology (Braak 0-III) and those with high levels of pathology (Braak IV-VI) after adjusting for the covariates of age, sex and batch effects. A positive Δ reflects increased levels of the modification with increasing pathology. All P values < 0.05 (and their corresponding Δ) are shown in bold.*

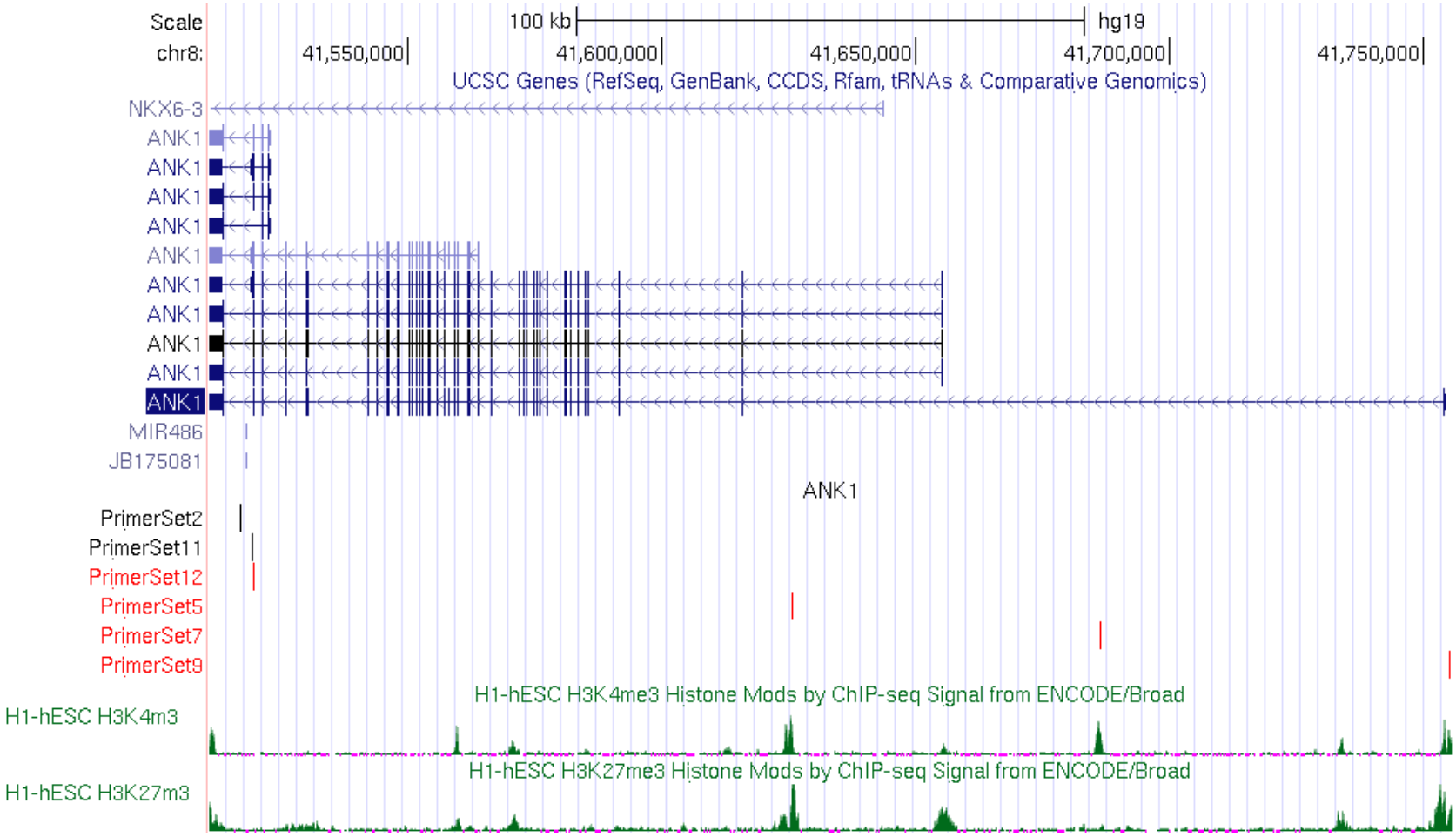


Figure 5.5; Visualisation of ChIP-qPCR primer sets used in the study. *Four primer sets showed decreased H3K4me3 in individuals with high levels of pathology compared to those with low levels of pathology. Displayed is the complete ANK1 gene with all 10 transcript variants (shown in purple), locations of primer sets that showed significantly decreased H3K4me3 levels with neuropathology (shown in red), locations of primer sets that showed no change in H3K4me3 levels with neuropathology (shown in black), as well as established H3K4me3 and H3K27me3 profiles in ESCs (shown in green).*

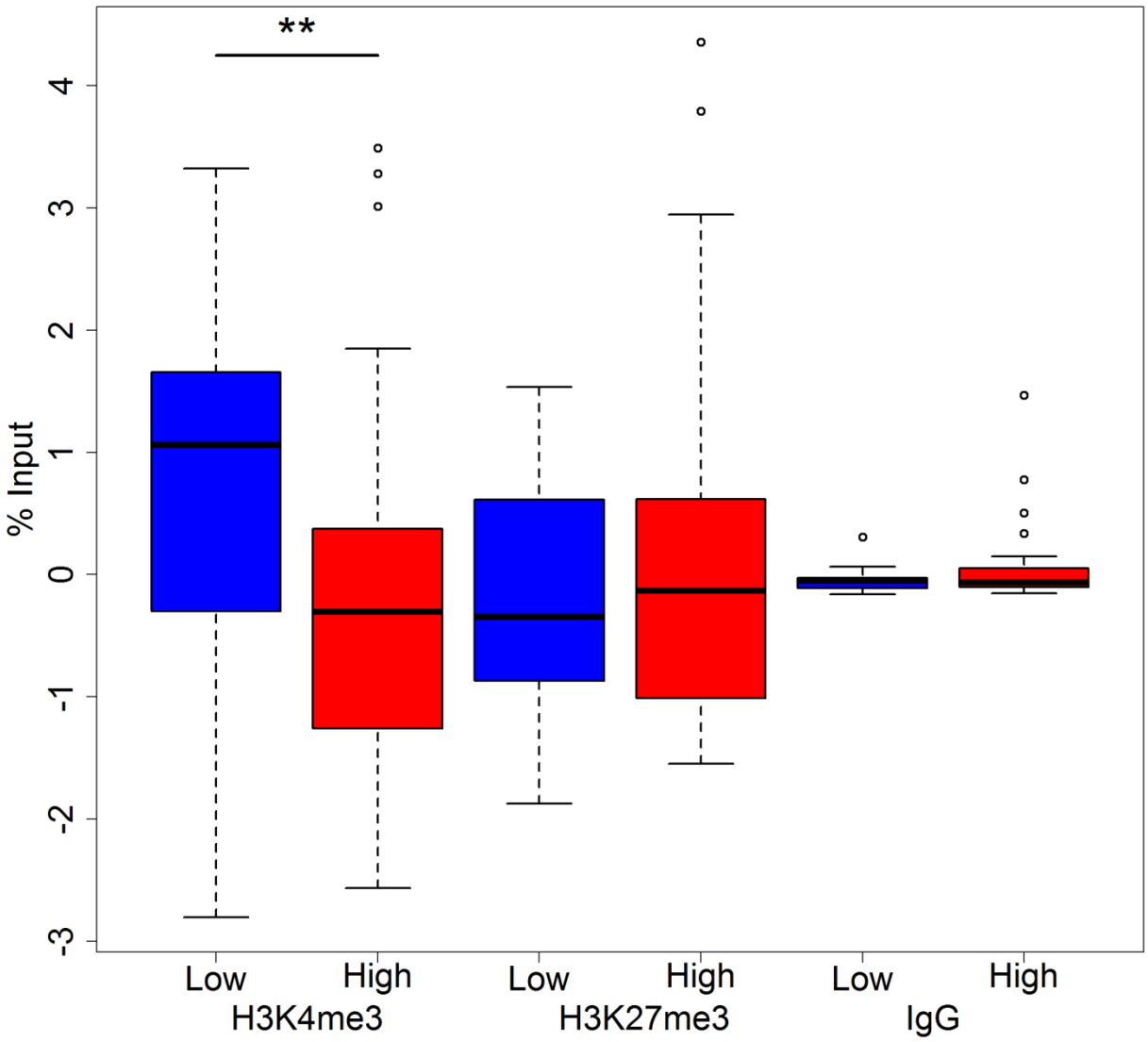


Figure 5.6; H3K4me3 is decreased across the ANK1 gene in individuals with high neuropathology burden. *When we averaged histone modification levels across the six regions of ANK1 we examined we identified a significant reduction in H3K4me3 levels in individuals with a high level of neuropathology burden, compared to those with a low burden (two sample t-test: t_{86} = -3.2, P = 0.002) meeting our Bonferroni threshold of P = 0.017, with no change in H3K27me3 or IgG control. Data is represented as a percentage relative to input. Box and whisker plot shows median, upper and lower quartiles, and maximum and minimum values (excluding outliers). Key: *P < 0.05, **P < 0.01, ***P < 0.005.*

5.4.2 H3K4me3 and H3K27me3 levels are correlated in post-mortem brain samples.

H3K4me3 and H3K27me3 histone modifications can have opposing effects on gene expression. However, both these modifications can simultaneously exist on the same histone tail (Bernstein et al. 2005). We were therefore interested to investigate whether there was a correlation of the levels of both modifications across the *ANK1* locus. As such, we found a positive correlation between H3K4me3 and H3K27me3 levels (r = 0.32, P = 0.002) (Figure 5.7).

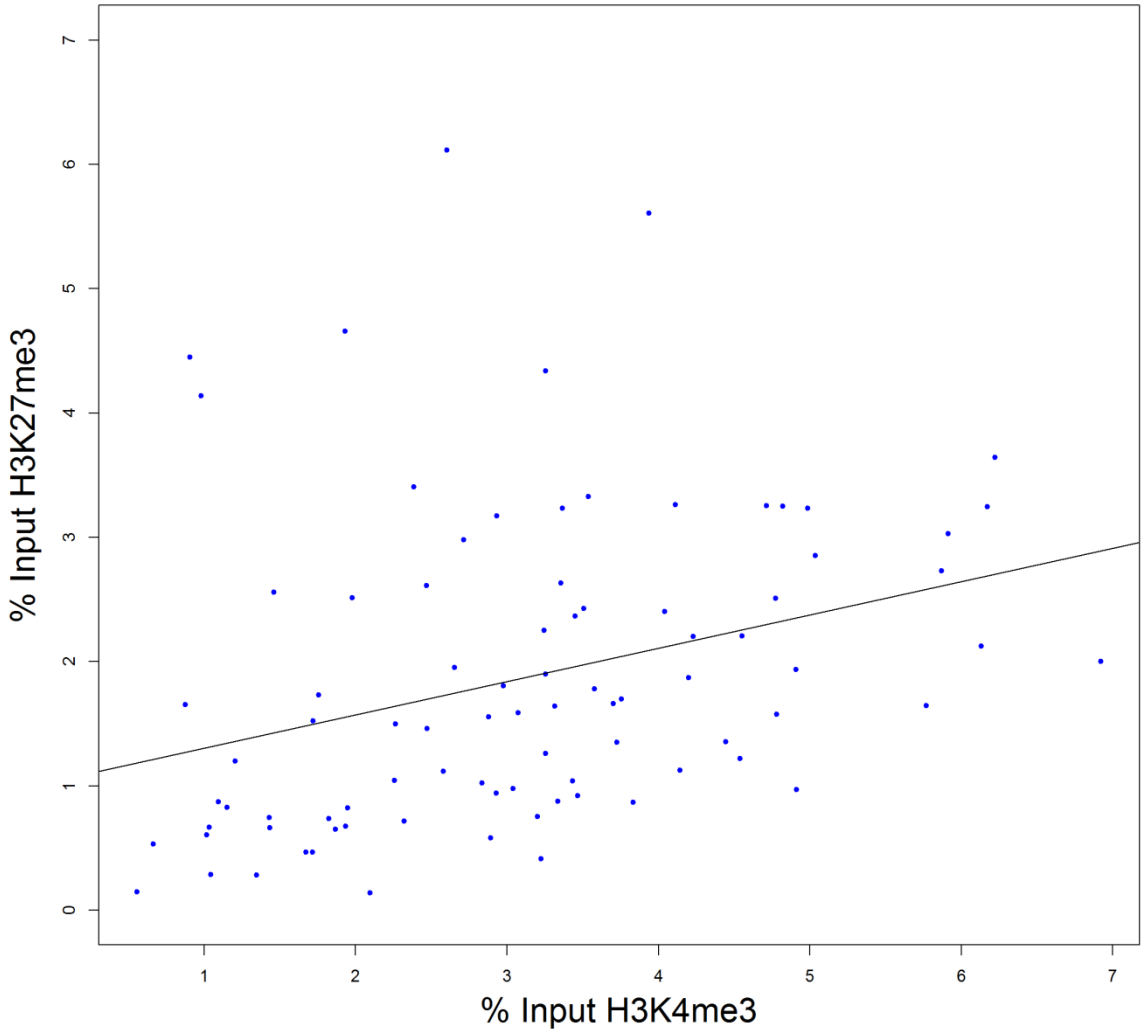


Figure 5.7; H3K4me3 Levels are significantly correlated with H3K27me3 levels in post-mortem brain samples. *H3K4me3 and H3K27me3 averaged across the six genomic regions interrogated are significantly correlated with each other (Pearson's correlation: t_{86} = 3.1, r = 0.32, P = 0.002).*

5.4.3 H3K4me3 levels are correlated with DNA modification levels in the *ANK1* gene

For a subset of 63 individuals used in this study we had already collected DNA methylomic and hydroxymethylomic data in the same samples (Chapter 3). Previous reports have demonstrated 5mC and H3K4me are (negatively) correlated with each other (Balasubramanian et al. 2012). To examine further the relationship between histone modifications and cytosine modifications, we first assessed the correlation between average H3K4me3 levels across all six regions assayed and the average 5mC or 5hmC level at 450K derived CpG sites across the *ANK1* gene (N = 107 probes). This showed that 5mC at 11% of CpG sites (12 / 107) within the *ANK1* gene was significantly correlated with H3K4me3 levels averaged across the six regions (Figure 5.8, Table 5.4). Similarly, 5hmC at 10% of CpG sites (11 / 107) was significantly correlated with H3K4me3 levels averaged across all six regions (Figure 5.9, Table 5.5). Five CpG sites were common to both analyses (cg08786207, cg08521995, cg13152952, cg09405790 and cg17256609), with an opposite direction of correlation seen for 5mC and 5hmC at each of these sites.

We then assessed the correlation between H3K4me3 levels in each individual region covered by the primers and 5mC or 5hmC levels at CpG sites covered by the 450K array in the same region (± 1kb) to see whether localised correlation existed. Three of the six regions showed significant correlation between H3K4me3 levels and 5mC levels (chr8:41516703-41516777 (± 1kb) (Primer Set 2): Figure 10A, r = -0.35, P = 0.005; chr8: 41519216-41519359 (± 1kb) (Primer Set 11): Figure 10B, r = -0.27, P = 0.039; and chr8:41519342-41519460 (± 1kb) (Primer Set 12): Figure 10C, r = -0.37, P = 0.033). No significant correlation was seen between H3K4me3 levels and 5hmC levels in the individual regions.

Finally, two specific 5mC sites in *ANK1*, cg11823178 (chr8:41519399) and cg05066959 (chr8:41519308), have been robustly shown to be hypermethylated with increasing AD neuropathology (De Jager et al. 2014, Lunnon et al. 2014, Smith et al. 2017) (Chapter 3 and 4). We were interested to investigate whether H3K4me3 levels were correlated with 5mC and 5hmC levels at these two specific sites, given that Primer Set 11 (chr8:41519216-41519359) covered cg05066959 and Primer Set 12

(chr8:41519342-41519460) covered cg11823178.This showed a significant negative correlation of H3K4me3 levels across chr8:41519342-41519460 (Primer Set 12) with 5mC levels at chr8:41519399 (cg11823178) (r = -0.36, P = 0.003) (Figure 5.11A). There was no correlation between H3K4me3 levels across that region and 5hmC levels (Figure 5.11B). There was no significant correlation between H3K4me3 levels across chr8:41519342-41519460 (Primer Set 11) with 5mC or 5hmC levels at chr8:41519308 (cg05066959).

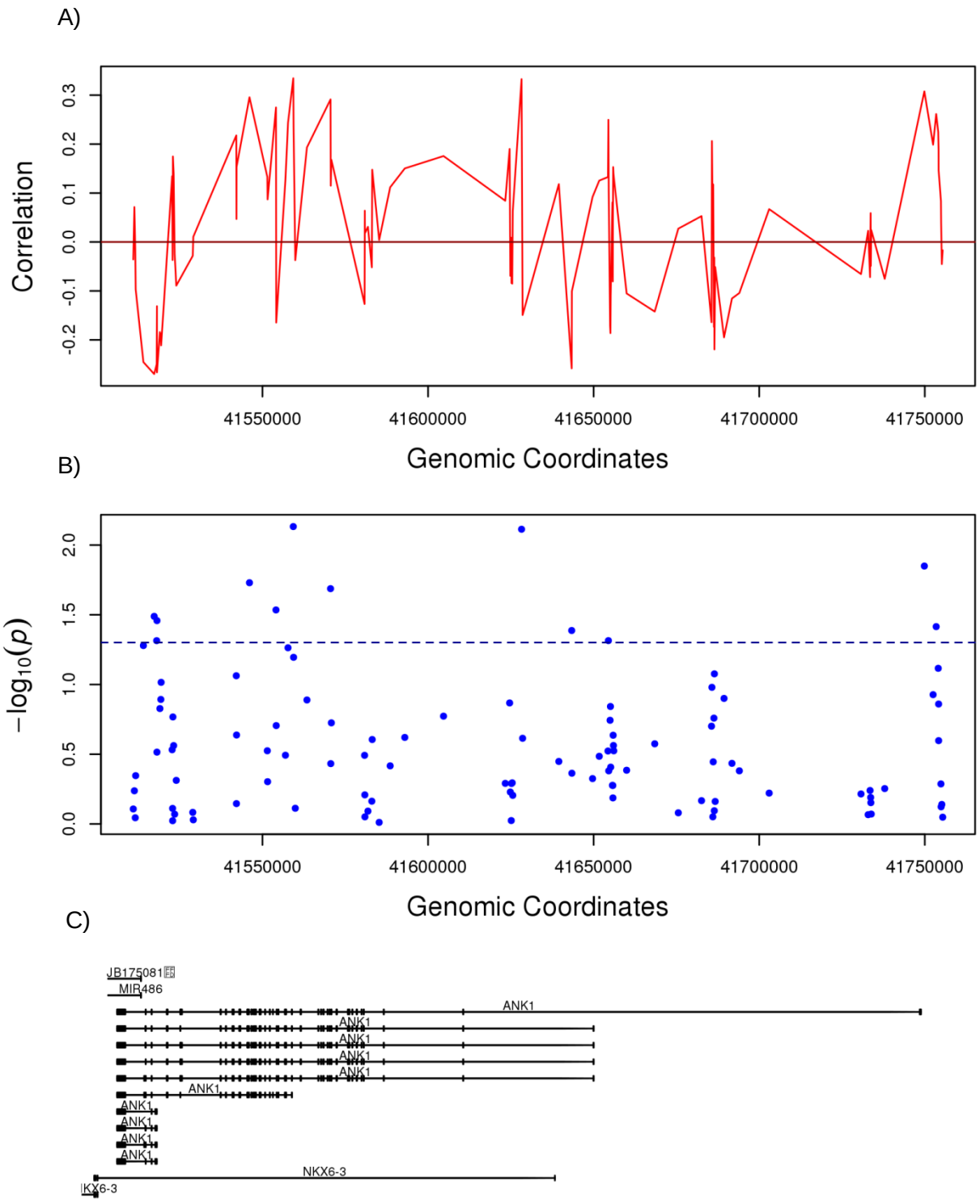


Figure 5.8; Average H3K4me3 levels are correlated with 5mC levels at specific CpG sites in ANK1. *(A) Pearson's correlation coefficient (r) between average H3K4me3 levels across the six regions tested and 5mC levels at specific CpG sites covered by the 450K array (N=107), in relation to ANK1 genomic location. (B) Manhattan plot the association between average H3K4me3 levels across the six regions tested and 5mC levels at specific CpG sites covered by the 450K array. The dashed line represents P = 0.05. (C) Schematic of the assessed genomic region, including all known isoforms of ANK1. No correlations met our multiple testing correction threshold (Bonferroni significance P < 4.67 x10^{-4}).*

Probe	Genomic Coordinates	*r*	*P*
cg12148515	chr8:41517301	-0.27	0.032
cg10794439	chr8:41518051	-0.25	0.048
cg10715527	chr8:41518139	-0.27	0.035
cg22532314	chr8:41546072	0.30	0.019
cg09576978	chr8:41554080	0.27	0.029
cg08786207	**chr8:41559303**	**0.33**	**0.007**
cg13152952	**chr8:41570554**	**0.29**	**0.021**
cg08521995	**chr8:41628237**	**0.33**	**0.008**
cg17256609	**chr8:41643375**	**-0.26**	**0.041**
cg08923054	chr8:41654455	0.25	0.048
cg22746182	chr8:41749867	0.31	0.014
cg09405790	**chr8:41753440**	**0.26**	**0.038**

Table 5.4; H3K4me3 levels are correlated with 5mC levels at specific CpG sites. *Shown for each 450K probe is genomic location (hg19), Pearson's correlation coefficient (r) and corresponding P value for the correlation between average H3K4me3 levels across the six regions tested and 5mC levels at individual CpG sites. Loci where average H3K4me3 levels were significantly correlated with both 5mC and 5hmC levels are shown in bold (N=5). Data is only shown for the 12 CpG sites that were correlated. No correlations met our multiple testing correction threshold (Bonferroni significance $P < 4.67 \times 10^{-4}$).*

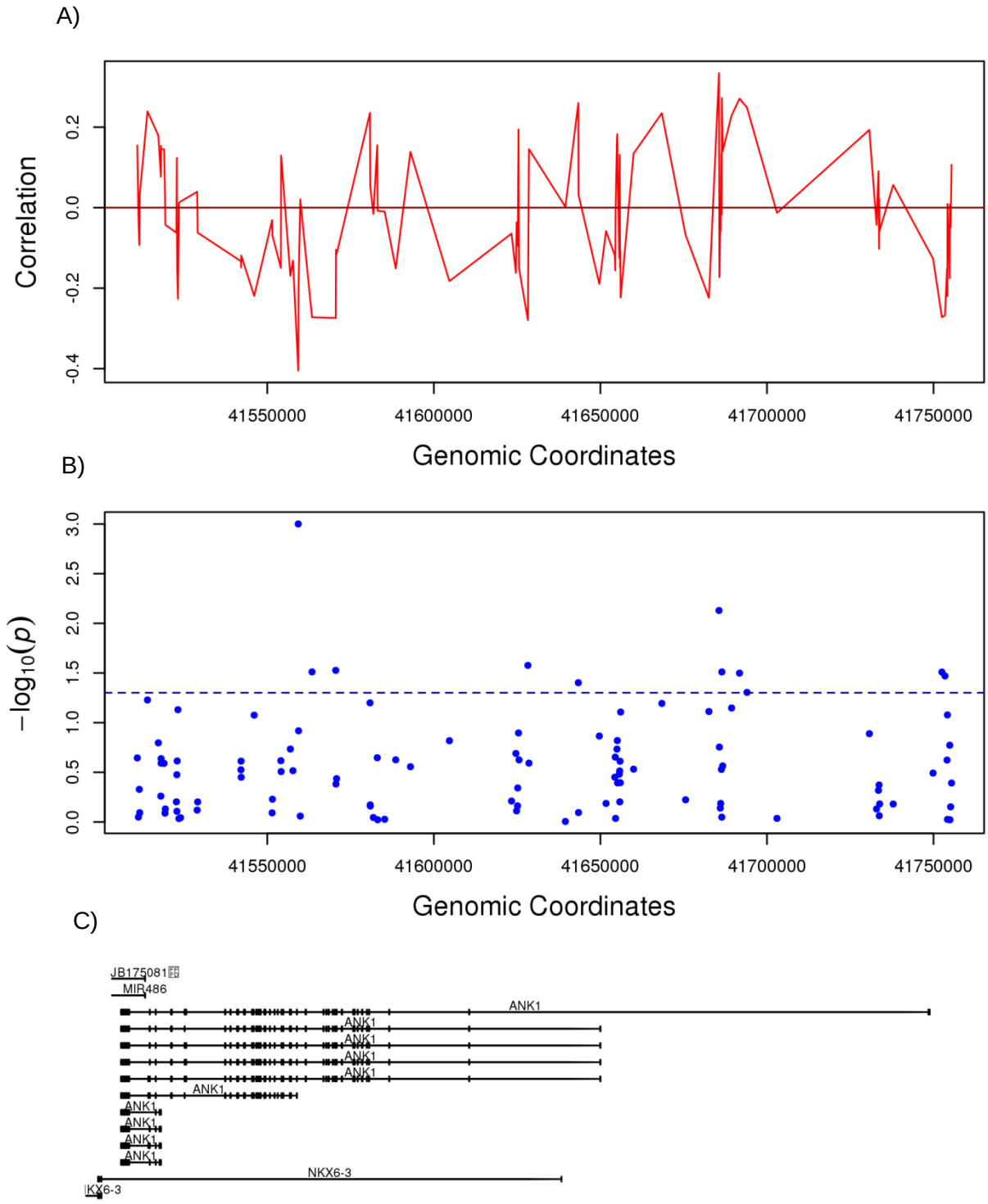


Figure 5.9; Average H3K4me3 levels are correlated with 5hmC levels at specific CpG sites in ANK1. *(A) Pearson's correlation coefficient (r) between average H3K4me3 levels across the six regions tested and 5hmC levels at specific CpG sites covered by the 450K array, with relation to ANK1 genomic location. (B) Manhattan plot showing the association between average H3K4me3 levels across the six regions tested and 5hmC levels at specific CpG sites covered by the 450K array. The dashed line represents P = 0.05. (C) Schematic of the assessed genomic region, including all known isoforms of ANK1. No correlations met our multiple testing correction threshold (Bonferroni significance $P < 4.67 \times 10^{-4}$).*

Probe	Genomic Coordinates	*r*	*P*
cg08786207	**chr8:41559303**	**-0.40**	**0.001**
cg07680965	chr8:41563429	-0.27	0.031
cg13152952	**chr8:41570554**	**-0.27**	**0.030**
cg08521995	**chr8:41628237**	**-0.28**	**0.027**
cg17256609	**chr8:41643375**	**0.26**	**0.040**
cg19997384	chr8:41685595	0.33	0.007
cg16129172	chr8:41686467	0.27	0.031
cg14354820	chr8:41691767	0.27	0.032
cg22845790	chr8:41694003	0.25	0.050
cg05977053	chr8:41752508	-0.27	0.031
cg09405790	**chr8:41753440**	**-0.27**	**0.034**

Table 5.5; H3K4me3 levels are correlated with 5hmC levels at specific CpG sites. *Shown for each probe is genomic location (hg19), Pearson's correlation coefficient (r) and corresponding P value for the correlation between average H3K4me3 levels across the six regions tested and 5hmC levels at individual CpG sites. Loci where average H3K4me3 levels were significantly correlated with both 5hmC and 5mC levels are shown in bold (N=5). Data is only shown for the 11 CpG sites that were correlated. No correlations met our multiple testing correction threshold (Bonferroni significance $P < 4.67 \times 10^{-4}$).*

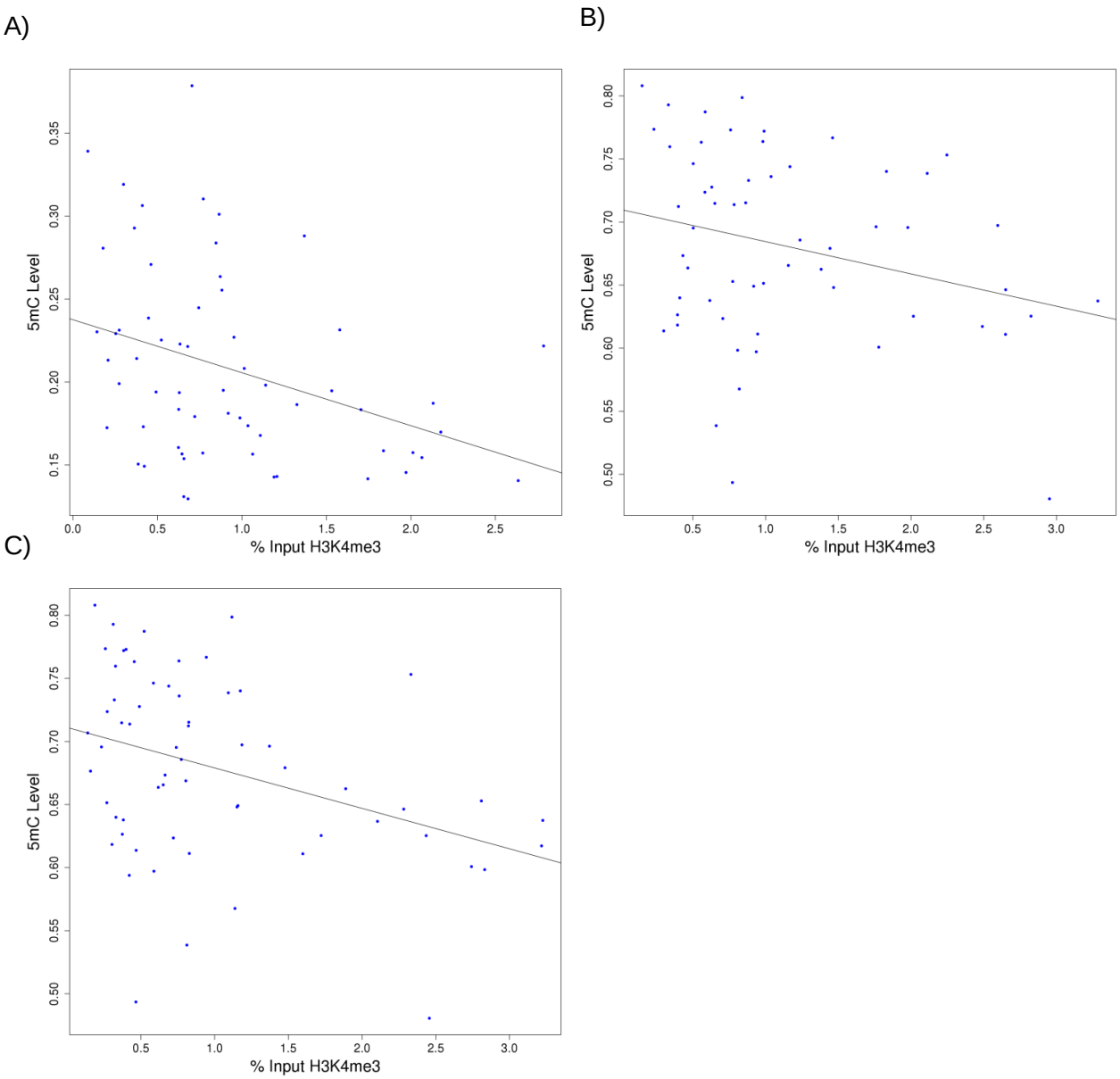


Figure 5.10; H3K4me3 levels in specific regions are correlated with 5mC levels in those regions. *H3K4me3 levels at three of the six genomic regions interrogated were significantly correlated (Pearson's correlation) with 5mC levels in those regions (± 1kb), including (A) chr8:41516703-41516777 (Primer Set 2): t_{60} = -2.9, r = -0.35, P = 0.005, (B) chr8: 41519216-41519359 (Primer Set 11): t_{56} = -2.1, r = -0.27, P = 0.039, and (C) chr8:41519342-41519460 (Primer Set 12): t_{61} = -3.0, r = -0.37, P = 0.033.*

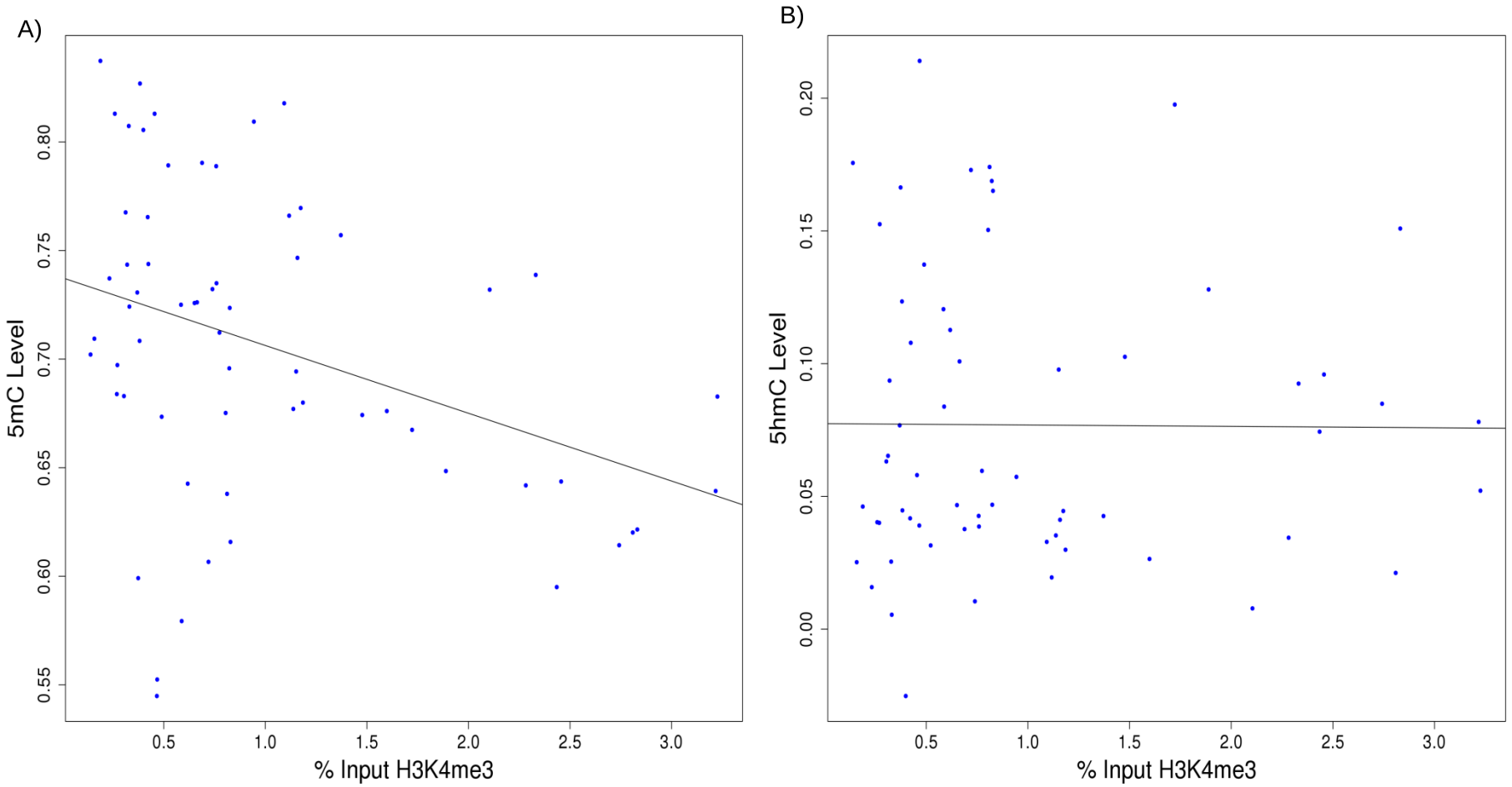


Figure 5.11; H3K4me3 levels are negatively correlated with 5mC levels at cg11823178. *H3K4me3 levels across chr8: 41519342-41519460 (Primer Set 12) are significantly correlated (Pearson's correlation) with (A) 5mC levels at chr8:41519399 (cg11823178: t_{61} = -3.1, r = -0.36, P = 0.003), (B) but not with 5hmC levels measured in matched tissue samples.*

5.4.4 H3K27me3 levels are correlated with DNA modification levels in the *ANK1* gene

Next, we were interested to investigate whether a similar relationship with DNA modifications was observed for H3K27me3. First, we assessed the correlation between H3K27me3 levels averaged across all six regions assayed and the average 5mC or 5hmC level at 450K derived CpG sites across the *ANK1* gene (N = 107 probes). This showed that 5mC levels at four of the CpG sites (Figure 5.12, Table 5.6) and 5hmC levels at seven of the CpG sites (Figure 5.13, Table 5.7) within the *ANK1* gene were significantly correlated with average H3K27me3 levels. Only one probe was common to both analyses (cg22845790), with an opposite direction of correlation observed for 5mC (r = -0.30, P = 0.018) and 5hmC (r = 0.34, P = 0.007).

When we looked within the individual regions assayed we observed no significant correlation between either 5mC or 5hmC levels with H3K27me3 levels in each individual region (± 1kb). Similarly, 5mC and 5hmC levels at the two methylation sites in *ANK1* that are known to be hypermethylated in AD (chr8:41519308 -cg05066959 and chr8:41519399 -cg11823178) were not correlated with H3K27me3 levels across chr8:41519216-41519359 (Primer Set 11) and chr8:41519342-41519460 (Primer Set 12), respectively.

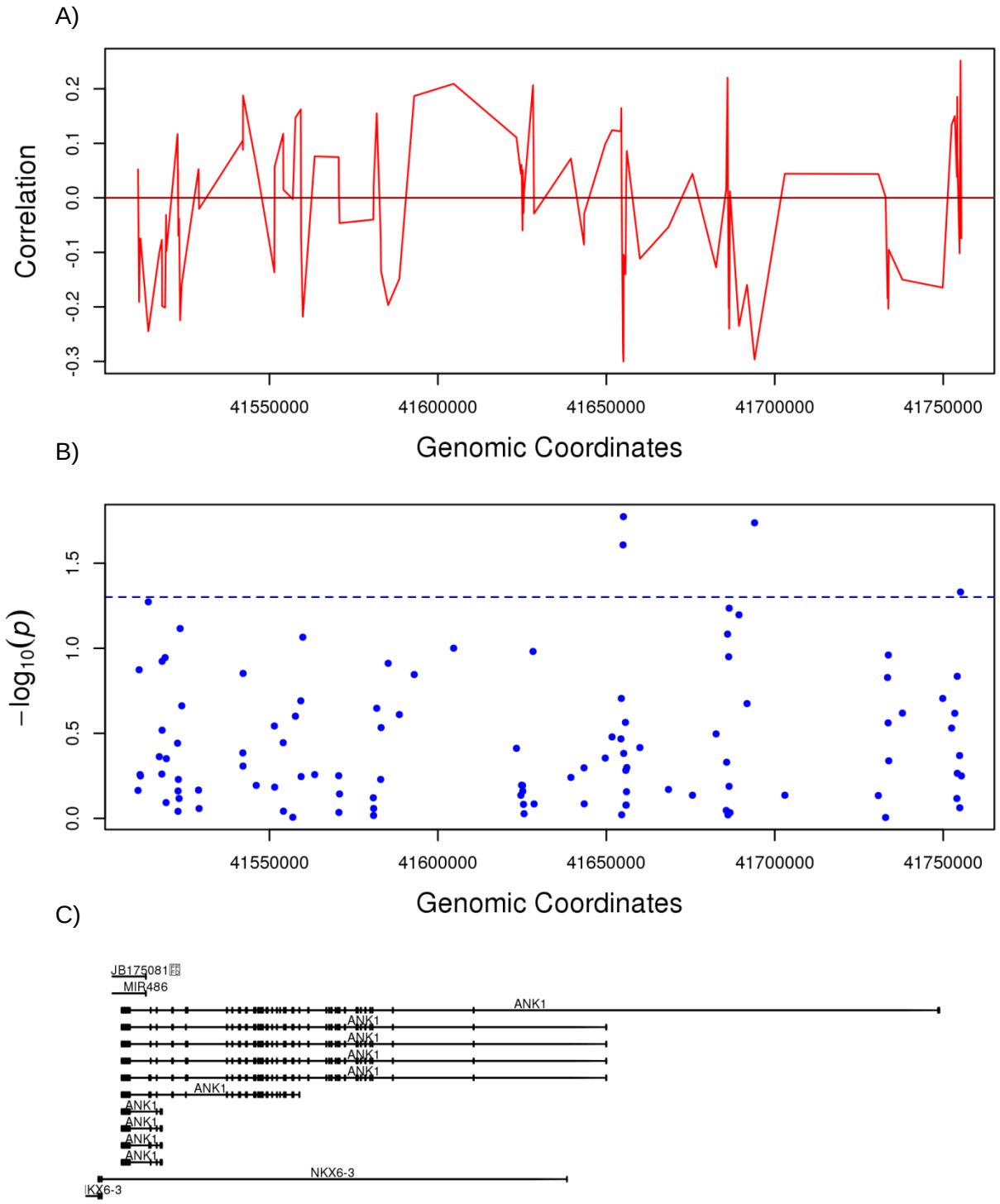


Figure 5.12; Average H3K27me3 levels are correlated with 5mC levels at specific CpG sites in ANK1. *(A) Pearson's correlation coefficient (r) between average H3K27me3 levels across the six regions tested and 5mC levels with relation to ANK1 genomic location. (B) Manhattan plot showing the association between average H3K27me3 levels across the six regions tested and 5mC levels at specific CpG sites covered by the 450K array. The dashed line represents P = 0.05. (C) Schematic of the assessed genomic region, including all known isoforms of ANK1. No correlations met our multiple testing correction threshold (Bonferroni significance P < 4.67 x10^{-4}).*

Probe	Genomic Coordinates	*r*	*P*
cg00176210	chr8:41654967	-0.28	0.025
cg26326633	chr8:41655078	-0.30	0.017
cg22845790	**chr8:41694003**	**-0.30**	**0.018**
cg01616178	chr8:41755140	0.25	0.047

Table 5.6: H3K27me3 levels are correlated with 5mC levels at specific CpG sites. *Shown for each probe is genomic location (hg19), Pearson's correlation coefficient (r) and corresponding P value for the correlation between average H3K27me3 level across the six regions tested and 5mC levels at individual CpG sites. Loci where average H3K27me3 levels were significantly correlated with both 5mC and 5hmC levels are shown in bold (N=1). Data is only shown for the 4 CpG sites that were correlated.*

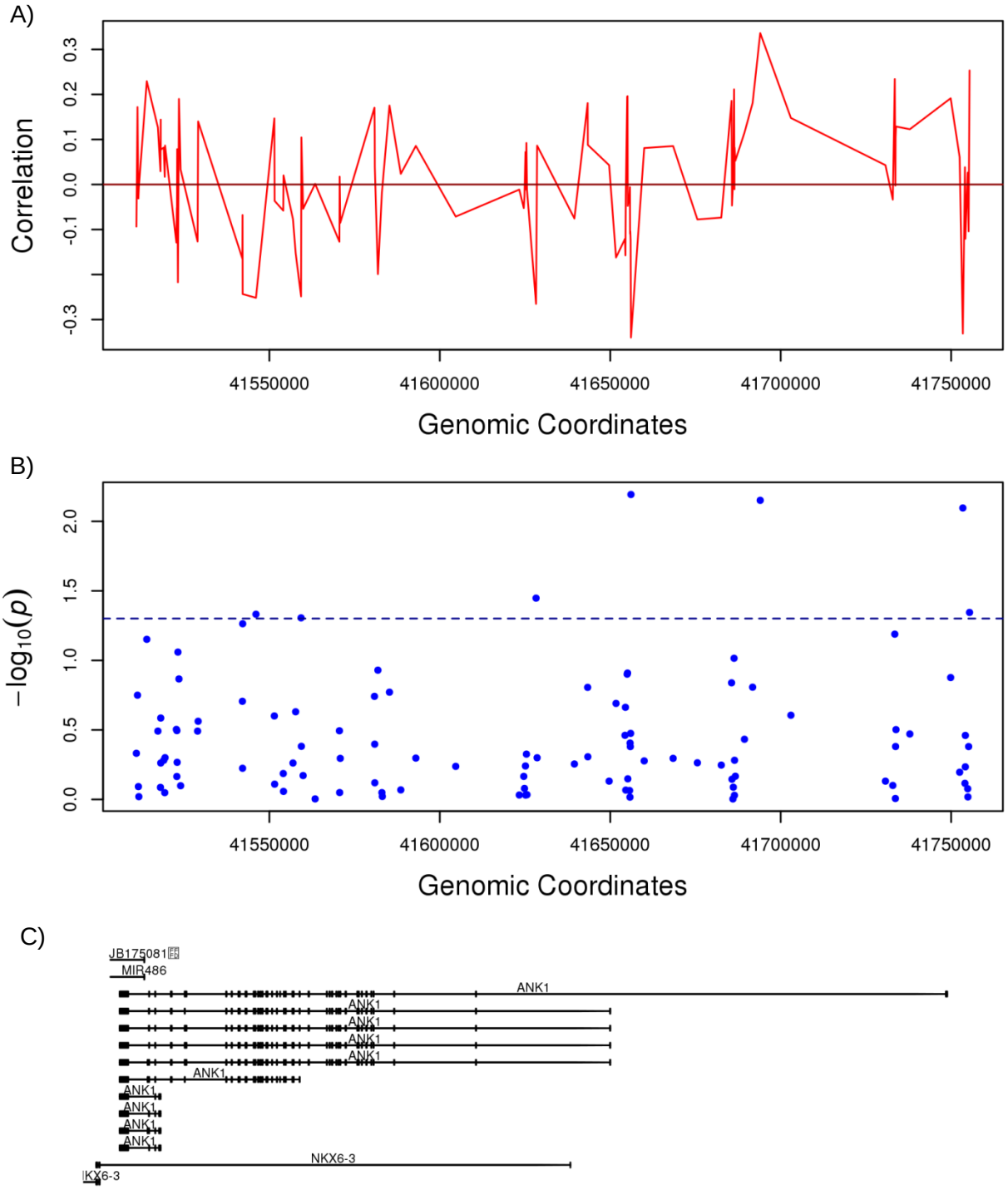


Figure 5.13; Average H3K27me3 levels are correlated with 5hmC levels at CpG sites in ANK1. *(A) Pearson's correlation coefficient (r) between average H3K27me3 levels across the six regions tested and 5hmC levels at specific CpG sites covered by the 450K array, with relation to ANK1 genomic location. (B) Manhattan plot showing the association between average H3K27me3 levels across the six regions tested and 5hmC levels at specific CpG sites covered by the 450K array. The dashed line represents P = 0.05. (C) Schematic of the assessed genomic region, including all known isoforms of ANK1. No correlations met our multiple testing correction threshold (Bonferroni significance P < 4.67 x10^{-4}).*

Probe	Genomic Coordinates	*r*	*P*
cg22532314	chr8:41546072	-0.25	0.047
cg08786207	chr8:41559303	-0.25	0.049
cg08521995	chr8:41628237	-0.27	0.036
cg26984624	chr8:41656079	-0.34	0.006
cg22845790	**chr8:41694003**	**0.34**	**0.007**
cg09405790	chr8:41753440	-0.33	0.008
cg19844326	chr8:41755409	0.25	0.045

Table 5.7; H3K27me3 levels are correlated with 5hmC levels at specific CpG sites. *Shown for each probe is genomic location (hg19), Pearson's correlation coefficient (r) and corresponding P value for the correlation between average H3K27me3 levels across the six regions tested and 5hmC levels at individual CpG sites. Loci where average H3K27me3 levels were significantly correlated with both 5hmC and 5mC levels are shown in bold (N=1). Data is only shown for the 7 CpG sites that were correlated. No correlations met our multiple testing correction threshold (Bonferroni significance $P < 4.67 \times 10^{-4}$).*

5.5 Discussion

This study has optimised the ChIP-qPCR protocol for use in post-mortem human brain tissue. We have measured two histone modifications: H3K4me3, a marker of gene expression, and H3K27me3, a marker of transcriptional repression, across six regions of the *ANK1* gene in human, post-mortem brain samples.

We first investigated whether *ANK1* H3K4me3 and H3K27me3 levels are altered in the EC of individuals with high levels of AD neuropathology, compared to those with low levels. This showed significantly less H3K4me3, with no change in H3K27me3 across the *ANK1* gene in individuals with higher levels of neuropathology (Braak stage IV and above). This pattern of histone changes would usually be associated with decreased gene expression and would suggest that *ANK1* gene expression would be reduced in individuals with AD neuropathology. However, a number of studies indicate that *ANK1* gene expression is increased in AD; Mastroeni and colleagues recently showed increased levels of *ANK1* gene expression in microglia in AD brain, although they showed no change in the astrocytes or neurones (Mastroeni et al. 2017). Whilst Lunnon et al. (2014) showed that only certain transcript variants have increased expression in AD, whilst other variants have no disease-associated changes. In our study we have analysed H3K4me3 and H3K27me3 in bulk EC tissue and it may be that different patterns of histone modifications are observed in the different cellular subtypes. In the future, further studies and analyses on cell sorted post-mortem tissue should be performed to determine if the changes we have observed are ubiquitous across cell types, or are specific to only certain cell types. Protocols for the separation of separate cellular subtypes from post-mortem brain tissue are currently being optimised, which would allow this type of analysis in the future (Guez-Barber et al. 2012, Krishnaswami et al. 2016).

We showed H3K4me3 and H3K27me3 levels across the *ANK1* gene are positively correlated in post-mortem EC brain tissue. H3K4me3 and H3K27me3 are associated with active and inactive gene transcription respectively (Harikumar et al. 2015). However, both these modifications have been shown to simultaneously exist on the same histone tail of genomes of ESCs (Bernstein et al. 2005). Nonetheless, these

bivalent marks are not unique to ESCs, and are in fact found in multiple cell types, including neuronal cells, at different stages of development (Pan et al. 2007, Golebiewska et al. 2009). The positive correlation of H3K4me3 and H3K27me3 in our study in post-mortem brain samples suggests a complex interplay of epigenetic mechanisms in the regulation of *ANK1* gene expression in the brain.

Cytosine modification (5mC and 5hmC) levels at a number of specific loci were found to correlate with average histone modification (H3K4me3 and H3K27me3) levels across the *ANK1* gene. H3K4me3 has been previously suggested to protect from 5mC at both CGIs and non-CGIs start sites (Balasubramanian et al. 2012). In our study, H3K4me3 levels were correlated with 5mC levels at 12 CpG sites, and 5hmC levels at 11 CpG sites, with five CpG sites common to both analyses and showing opposite directions of correlations for H3K4me3 with 5mC and 5hmC. H3K27me3 levels were correlated with 5mC levels at four CpG sites and 5hmC levels at seven CpG sites, with one locus common to both analyses and showing an opposite direction of correlation for H3K27me3 with 5mC and 5hmC. Interestingly, 5hmC levels at four CpG sites (cg08786207, cg08521995, cg22845790 and cg090450790) were correlated with both H3K4me3 and H3K27me3 levels. Three of the probes showed an opposite direction of correlation with H3K4me3 to H3K27me3, as would be expected, given that they are markers of gene expression and repression, respectively. However, 5hmC levels at one locus (cg22845790) showed a positive correlation with both H3K4me3 and H3K27me3 levels. It was previously thought that 5hmC was a transient epigenetic mark, and did not play a role in transcriptional regulation. 5hmC has been previously identified at high levels in the developing (Spiers et al. 2017) and adult brain (Munzel et al. 2010), particularly in neurones (Kriaucionis et al. 2009) and, as such, may represent an important epigenetic mark to profile in the context of neurodegenerative diseases. One study has shown that 5hmC levels in ESCs are more strongly correlated with levels of other histone modifications (i.e. H3K4me1, H3K4me2, H3K18ac, H3K27ac, H4K5ac), than they are with H3K4me3 or H3K27me3 (Szulwach et al. 2011). This suggests that there is likely a complex relationship between 5hmC levels and histone modification levels in *ANK1*.

5mC levels at two CpG sites (cg11823178: chr8:41519399 and chr8:41519308: cg05066959) have been previously shown to be hypermethylated in AD in numerous published studies (De Jager et al. 2014, Lunnon et al. 2014, Smith et al. 2017) and in Chapters 3 and 4 of this thesis. Interestingly, we found H3K4me3 levels across chr8:41519342-41519460 (Primer Set 12) to be significantly decreased in individuals with high neuropathology and to be negatively correlated with 5mC levels at chr8:41519399 (cg11823178). It has previously been suggested that H3K4 modifications are a preventative mechanism for *de novo* promoter cytosine methylation (Weber et al. 2007). A proposed mechanism for this stems from the discovery that Dnmt3L, a Dnmt3 associated protein required for efficient *de novo* cytosine methylation, contains a domain that specifically interacts with unmodified H3K4, the binding of which is inhibited by methylation of the H3 tail (Ooi et al. 2007). Furthermore mutation of this same domain in Dnmt3L leads to reduced 5mC levels (Rose et al. 2014). These observations suggest that H3K4 methylation may play a role in blocking *de novo* 5mC at some genomic loci (Ooi et al. 2007, Otani et al. 2009, Rose et al. 2014). It is therefore a possibility that the decreased H3K4me3 levels in this region are facilitating the increased 5mC at chr8:41519399 (cg11823178) in AD. However, as we did not see any association of H3K4me3 levels across chr8:41519216-41519359 (Primer Set 11) with 5mC at chr8:41519308 (cg05066959), it is difficult to draw any firm conclusions as to whether hypermethylation at chr8:41519399 in AD is caused by the reduction in H3K4me3 levels, or vice versa. In the future it will be of interest to quantify the levels of other histone modifications in the *ANK1* gene. Furthermore, these levels should be integrated with levels of 5mC, 5hmC and miRNAs to build up a complete picture of the epigenetic landscape of the *ANK1* gene in AD. It is not possible to determine the order of events leading up to the changes in H3K4me3 in AD that have been reported here. It is feasible, given the highly dynamic nature of histone modifications, these changes are as a result of, rather than the cause of, AD pathology. In the future it will therefore be of interest to study the temporal pattern and order of epigenetic changes in AD.

H3K4me3 has been previously shown (with H3K9me1) to be predictive of expression levels in low CpG content promoters (LCPs), whilst H3K27ac and H4K20me1 are predictive of high CpG content promoters (HCPs) (Karlić et al. 2010). To date two

EWAS of histone modifications in AD have been performed in post-mortem brain tissue. The first, quantified H3K27ac across the genome in post-mortem EC tissue samples, identifying widespread AD-associated variation (Marzi et al. 2017). Interestingly, differentially acetylated regions were identified in the vicinity of genes previously implicated in AD neuropathology, for example *APP*, *PSEN1*, *PSEN2* and *MAPT*. The second EWAS quantified H3K9 acetylation (H3K9ac) in post-mortem prefrontal cortex, in a large number of samples (N = 669). This showed that tau protein burden had a broad effect on the epigenome, affecting over 5,500 H3K9ac domains. In our study we have covered areas of known H3K4me3 and H3K27me3 peaks in the brain using qPCR assays. However, this method still only profiles a small amount of a relatively large (244 kb) gene. As such, in the future an H3K4me3 and H3K27me3 EWAS in AD using ChIP-Seq would be the optimal next step to completely characterise the levels of these histone modifications in AD brain.

5.6 Conclusions

This study is the first to interrogate H3K4me3 and H3K27me3 modification profiles across the *ANK1* gene in AD brain tissue. Overall, these results suggest that H3K4me3 levels are reduced in multiple regions in AD EC, and are correlated with changes in 5mC and 5hmC.

CHAPTER 6: DISCUSSION AND FUTURE PERSPECTIVES

6.1 Introduction

The primary aim of this thesis was to expand on the existing knowledge about the epigenetic landscape in AD, with a particular focus on DNA and histone modifications in the vicinity of the *ANK1* gene. This discussion will summarise the key findings from each chapter of this thesis and the limitations present within each study. Subsequently, it will discuss and attempt to address the limitations when studying epigenetic mechanisms in AD. I will emphasise how the research presented in this thesis has tackled some of these issues and outline future perspectives for the field, which would acknowledging issues that I was unable to address in this thesis.

6.2 Key findings from this thesis

6.2.1 Chapter 3: DNA methylomic and hydroxymethylomic signatures of AD brain

The aim of this project was to utilise genome-wide profiling technology to quantify DNA methylomic and DNA hydroxymethylomic variation simultaneously in human post-mortem brain samples with varying degrees of AD pathology. This study represented the first EWAS to independently measure 5hmC and 5mC in AD brain.

This study generated a large amount of data and, as such, it was possible to address multiple research questions. My first analysis focussed on global levels of 5mC and 5hmC in disease, given that previous attempts to characterise global 5hmC in AD have been performed using immunocytochemistry and therefore provided global levels of the modification only (Chouliaras et al. 2013, Condliffe et al. 2014, Coppieters et al. 2014, Lashley et al. 2015). To this end my approach was to average 5mC and 5hmC levels across all probes that passed QC to compare how global levels of these modifications are altered in AD using Braak stage as a measure of disease severity. A global decrease in 5hmC and an increase in 5mC in the EC with respect to Braak stage was observed, whilst no significant Braak stage associated alterations were seen in the CER. Although, Condliffe et al. (2014) found similar patterns in the EC, other studies have focused on different brain regions with varying results (Chouliaras et al. 2013, Coppieters et al. 2014, Lashley et al. 2015). My results suggest that global changes in the levels of cytosine modifications occur in a tissue specific manner in AD. However, it must be noted that only ~2% of potential modification sites in the human genome are covered by the 450K array and thus the global levels I found to be altered in AD in Chapter 3 are not necessarily representative of all sites in the human genome.

My next analysis focussed on identifying DMoPs, DMPs and DHPs specifically altered with respect to AD pathology (Braak stage). My analyses demonstrated highly reproducible alterations in Braak-associated DNA modification changes in the EC; I saw that the effect sizes of DMoPs were highly correlated between my study and a previous publication. Interestingly, there was no correlation of the effect sizes

of DMoPs in the CER between studies, suggesting that the brain regions which are relatively protected from neurodegeneration do not display consistent disease-associated epigenetic changes in total DNA modifications. My next analysis identified 14 CpG sites, which showed significant Braak-associated alterations for more than one analysis (DMoPs, DMPs or DHPs). Of these, 13 showed no nominally significant difference in total DNA modifications, but showed a significant difference in both 5mC and 5hmC, this included the *APP* gene known to cause FAD. In each instance the levels of both modifications were highly negatively correlated with each other and showed a significant interaction between the two modifications. In the future, it will be important to validate these loci in other sample cohorts and to determine the effect these changes have on gene expression, as unlike 5mC which is commonly associated with gene silencing, 5hmC has been shown in genomic regions which in turn positively correlate with gene expression (Pastor et al. 2011, Song et al. 2011, Wu et al. 2011, Xu et al. 2011). Interestingly, my analyses showed that neuropathology-associated changes in 5mC and 5hmC were most pronounced at CGIs of proximal promotors, which are regions key for the binding of transcription factors leading to gene expression (Spitz et al. 2012).

I identified 72 Braak-associated DMPs and 57 DHPs that reached our significance threshold, of which 3 DMPs reached experiment-wide significance. A pathway analysis of the significant Braak-associated DMPs and DHPs highlighted common pathways to both analyses, which were associated with the axon, with DHPs being specifically enriched in these biological pathways associated with the synapse. Dysregulation of these pathways could be a driving factor in the pathology of AD, as the disease is characterised by synaptic loss and neuronal death (Whitehouse et al. 1982). My final analysis focussed on identifying Braak-associated DMRs and DHRs. I did not identify any DHRs, which was not surprising given that 5hmC has been reported to be enriched in areas of low CG density (Lunnon et al. 2016). I did identify four DMRs. Of particular interest was a DMR within the *ANK1* gene, as two previous EWAS of AD have reported a DMoR in this region (De Jager et al. 2014, Lunnon et al. 2014). Using pyrosequencing I replicated this DMR using pyrosequencing to assay eight neighbouring CpG sites, including the two covered by the array. Of the eight CpG sites assessed, seven were characterised by significant AD-associated hypermethylation, and four were characterised by significant AD-associated

hypohydroxymethylation. My study demonstrated that previous estimates of *ANK1* hypermethylation from EWAS using BS-treated DNA were underestimates (De Jager et al. 2014, Lunnon et al. 2014), due to confounding hypohydroxymethylation (Smith et al. 2016).

The 450K array paved the way for EWAS as it is relatively cost-effective, high throughput and numerous analysis pipelines have been developed. However, it is not without limitations. Many of the array's design faults, including the presence of SNP and cross-hybridising probes, were addressed in the analysis pipeline (Section 2.8). However, as the array only covers approximately 485,000 CpG positions of the 28 million in the human genome (Lovkvist et al. 2016), over 98% of potential modification sites are not assayed, meaning the likely loss of some disease specific variation. Another caveat of all epigenetic studies of AD brain to date, including the study in Chapter 3, is the issue of cellular diversity in the brain, with this issue further confounded in AD, where the abundance of specific cell types changes in the disease. A further limitation, in terms of the ability to translate these findings to novel pharmacological targets, is that it is not possible to determine whether these cytosine modification changes are the cause or consequence of disease.

In summary, I quantified 5hmC and 5mC levels simultaneously in post-mortem AD brain in Chapter 3. In doing so we identified 13 novel epigenetically altered loci that showed differential 5mC and 5hmC levels in disease and highlighted that the previous estimates of differential *ANK1* 5mC in AD brain were underestimates.

6.2.2 Chapter 4: *ANK1* DNA modification in different neurodegenerative diseases

This study aimed to determine DNA modification levels in a number of different dementias across a region of the *ANK1* gene which was previously reported to display hypermethylation and hypohydroxymethylation in AD (Chapter 3). Using BS pyrosequencing, *ANK1* 5mC was assessed in brain samples from donors with AD, DLB, VaD, HD, PD and non-demented elderly controls, across a number of different brain regions that show disease-specific pathology.

Initially I replicated the *ANK1* Braak-associated DMoR I reported in the EC in Chapter 3 in additional AD samples. We showed DNA hypermethylation in both the EC and STG, replicating previous EWAS and pyrosequencing studies. Interestingly, we also observed significant *ANK1* hypermethylation in the CER, a brain region that previous EWAS had shown no significant *ANK1* modification change ((Lunnon et al 2014, Chapter 3, (Smith 2018)). On closer examination it became apparent that the sites of AD-associated hypermethylation in the CER were loci that were not covered on the 450K array and therefore would have missed detection in the previous EWAS. It was also interesting to note that the increase in 5mC difference in disease correlated with tissue specific levels of AD pathology, with the greatest 5mC difference seen in the EC, followed by the STG and then the CER, a region with very little AD pathology. When we assayed the *ANK1* region in other neurodegenerative diseases, it became clear that *ANK1* hypermethylation was only present in DLB and VaD EC in individuals with co-existing AD pathology and was also not seen in any of the other tissues tested. However, I did observe disease-associated *ANK1* DNA hypermethylation in the EC in individuals with HD and PD cases (without any other diagnosis), but observed no difference in the regions of primary pathology for these diseases, namely the STR and SN, respectively. This suggests that *ANK1* DNA hypermethylation is possibly present in a certain population of cells in the EC, which are altered in AD, HD and PD.

Similarly to the previous study (Chapter 3) one potential caveat of this study is that "bulk" tissue has been used, and therefore it is not possible to determine which cell type(s) are driving the DNA hypermethylation seen in *ANK1* in disease. Furthermore, opposing changes in different cell types could be confounding these results. Unlike Chapter 3, Chapter 4 utilised standard BS conversion, therefore the levels of 5mC reported are actually the sum of 5mC and 5hmC modifications. From previous results this would suggest that the levels reported are in fact underestimates of 5mC. Another caveat to my study is that we have only analysed 5mC across eight CpG sites in a 118bp region of the *ANK1* gene, which in its entirety is 244kb. Finally, and similarly to Chapter 3, at present it is not possible to determine whether the reported *ANK1* hypermethylation across the region in AD, HD and PD in the EC represents a cause or a consequence of the disease process.

In summary, I successfully determined total DNA modification levels across the previously identified *ANK1* locus in several neurological diseases including AD, HD, PD, DLB and VaD. This study showed differential levels of DNA modifications in the EC of AD, HD and PD across the *ANK1* locus but not in the other neurodegenerative diseases tested.

6.2.3 Chapter 5: Exploring histone modifications at the *ANK1* locus

The epigenetic studies of *ANK1* in AD in this thesis have focussed on levels of DNA modifications. To date, no published studies have examined the histone modification profile of this gene in relation to AD pathology. Therefore the aim of this chapter was to use ChIP-qPCR to quantify H3K4me3 and H3K27me3 modification levels at specific regions across the *ANK1* gene in AD.

This study revealed that there is a decrease in the level of the H3K4me3 histone modification at specific sites across the *ANK1* gene, but no significant change in H3K27me3. H3K4me3 is a modification usually associated with gene expression (Azuara et al. 2006, Kouzarides 2007, Mattout et al. 2010); therefore these results would suggest that *ANK1* expression would be decreased in the EC in AD brain. In my study H3K4me3 and H3K27me3 levels were significantly correlated, which concurs with previous studies that have shown H3K4me3 and H3K27me3 can occur simultaneously despite their conflicting functions (Bernstein et al. 2006, Harikumar et al. 2015). Bernstein and colleagues have shown that both these modifications exist on the same histone tail of genomes of ESCs, leading to highly dynamic changes in the expression and repression of the candidate gene (Bernstein et al. 2006). Similarly, I hypothesise that a similar mechanism could be regulating *ANK1* gene expression in AD brain. Previous studies have shown that there is cross talk between DNA and histone modifications (Kondo 2009). In Chapter 5 I showed that both total H3K4me3 and H3K27me3 levels correlated with 5mC and 5hmC levels at specific loci across the *ANK1* gene, suggesting a complex interplay of epigenetic regulation. In addition, I showed that H3K4me3 levels are negatively correlated with 5mC but not 5hmC, at chr8:41,519,399 (cg11823178), a loci which was previously reported to be highly significantly differentially methylated in Chapter 3 and in

multiple published studies (De Jager et al. 2014, Lunnon et al. 2014, Smith et al. 2018).

Akin to previous chapters one caveat of this study is that it utilised “bulk” brain tissue from the EC, meaning that I was unable to account for cellular heterogeneity. Second, histone modifications are very dynamic changes to the DNA, more so than cytosine modifications (Barth et al. 2010), and as I have assessed DNA in post-mortem tissue, it is not possible to determine the temporal order of events leading up to the changes in H3K4me3 that were reported in Chapter 5. Finally, despite my best efforts to cover genomic areas of brain specific histone modification peaks within the qPCR assays, this method still only profiles a small amount of what is a large and complex gene. Other areas of the gene that I was unable to profile in the current study will require further investigation in AD.

In summary, in Chapter 5 I profiled H3K4me3 and H3K27me3 profiles in the *ANK1* gene in AD EC, highlighting disease-associated decreases in H3K4me3, which are negatively correlated with 5mC levels at chr8:41,519,399 (cg11823178).

6.2.4 *ANK1* functional hypothesis

I have shown *ANK1* epigenetic differences exist in AD brain (Table 6.1). Functionally I hypothesise that *ANK1* epigenetic differences in disease are resulting in either a change in total *ANK1* expression or changing the expression level of a specific *ANK1* isoform. This change in *ANK1* abundance in the brain would then impact on the physiology of microglial cells. I hypothesise that the binding of the cytoskeleton to the plasma membrane of microglia is altered due to *ANK1* epigenetic changes in disease, in a similar manner to erythrocytes in HS. This, in turn would contribute to the “priming” of microglia within the brain, causing them to adopt a more aggressive inflammatory phenotype when exposed to a triggering event, leading to neuronal cell death and AD (Figure 6.1). This proposed hypothesis fits with the well described inflammatory pathway (Section 1.1.1.2).

Marker	Method	*ANK1* position	EC	STG	CER
DNA Methylation (Chapter 3 & 4)	BS Array	chr8:41518051	↑	-	-
	BS & OxBS Pyrosequencing	chr8:41519302	↑	↑	↑
		chr8:41519304	↑	↑	↑
	BS Array, BS & OxBS Pyrosequencing	chr8:41519308	↑	↑	-
	BS & OxBS Pyrosequencing	chr8:41519348	↑	↑	↑
	OxBS Array, BS & OxBS Pyrosequencing	chr8:41519399	↑	-	-
	BS & OxBS Pyrosequencing	chr8:41519411	↑	↑	↑
		chr8:41519417	↑	↑	↑
		chr8:41519420	↑	↑	↑
DNA Hydroxymethylation (Chapter 3)	OxBS Pyrosequencing	chr8:41519302	↓	-	-
		chr8:41519304	-	-	-
		chr8:41519308	-	-	-
		chr8:41519348	↓	-	-
		chr8:41519399	-	-	-
		chr8:41519411	-	-	-
		chr8:41519417	↓	-	-
		chr8:41519420	↓	-	-
H3K4me3 (Chapter 5)	ChIP-qPCR	chr8:41625416-41625491	↓	-	-
		chr8:41686181-41686328	↓	-	-
		chr8:41754877-41755012	↓	-	-
		chr8:41519342-41519460	↓	-	-

Table 6.1: Significant neuropathology-associated epigenetically modified findings in the ANK1 gene. *Epigenetic marker, profiling technology, chromosomal location (hg19) and whether a significantly positive (↑) or negative (↓) association was identified in the ANK1 gene in each brain region studied (entorhinal cortex (EC), Superior temporal gyrus (STG) and cerebellum (CER)).*

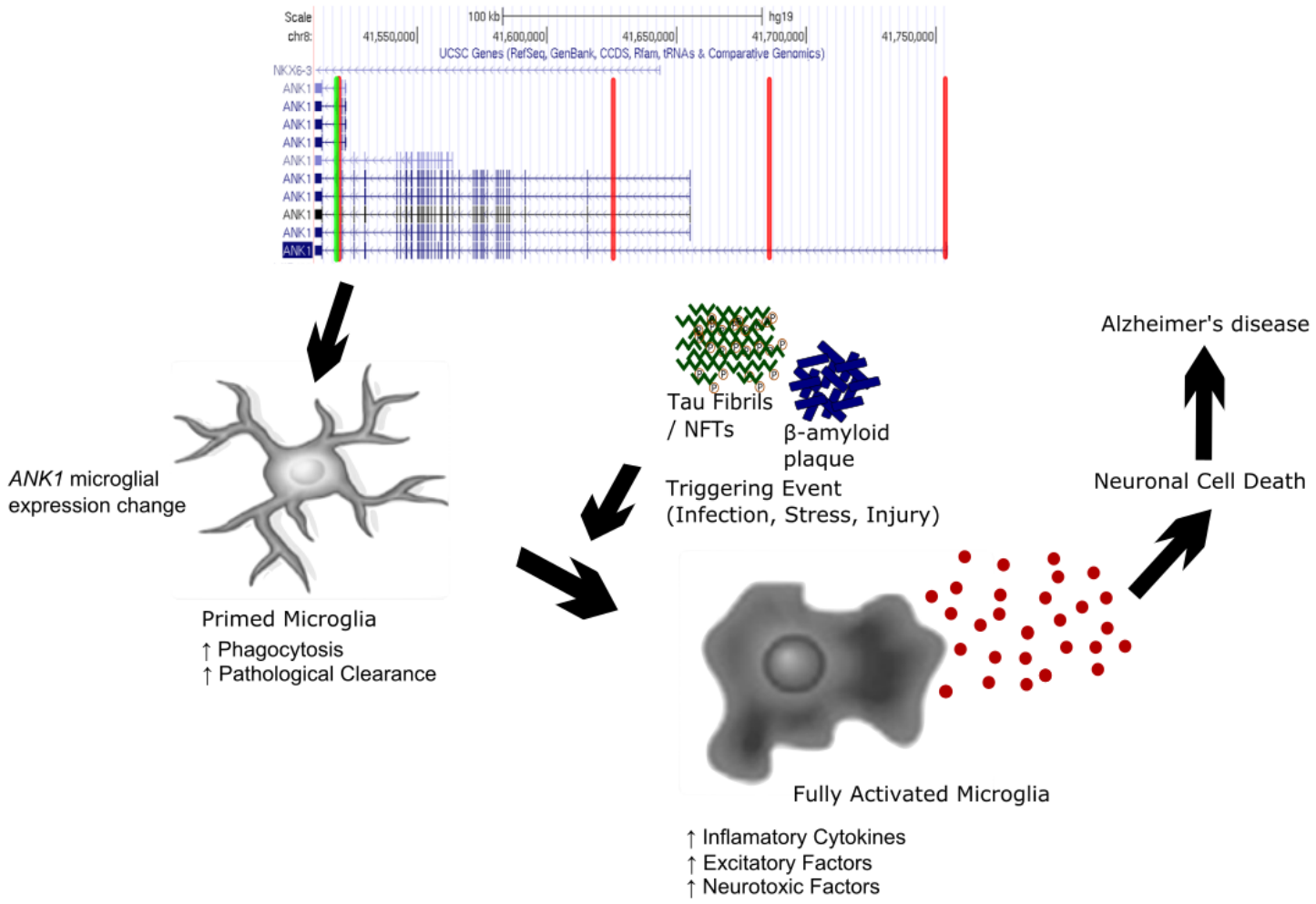


Figure 6.1; Proposed function of ANK1 epigenetic differences in AD brain. *ANK1 epigenetic alterations, increased 5mC and/or decreased H3K4me3 levels lead to a change in microglial ANK1 expression. This enables conformational changes, creating "primed" microglia that go on to cause AD through via the inflammatory hypothesis. Green and red bars represent DMR and H3K4me3 differences identified in this thesis.*

6.3 General Discussion and Future Perspectives

As detailed above, there are a number of limitations associated with studying epigenetic mechanisms in AD. There are however emerging methodologies that are being developed that address some of these issues. This next section will discuss the issues that have arisen in the studies presented in this thesis, with a particular emphasis on how I have attempted to consider these in the studies I have presented, as well as outlining future perspectives for the field that could address issues that I was unable to address at the present time.

6.3.1 Cellular Heterogeneity

Epigenetic profiles are in their nature cell specific, with individual cellular populations having a distinct epigenetic profile; therefore it is likely that disease related epigenetic changes could be affecting just one or two cellular populations, rather than every cell type in a tissue. However, every study reported in this thesis has utilised bulk tissue. Therefore epigenetic changes that have been reported in this thesis are representative of total cellular populations, potentially masking more biologically relevant singular population differences in AD. Chapter 3 attempted to address this issue using the bioinformatic CETS correction method, which corrects for neuron/glia proportions (Guintivano et al. 2013). Adding this value as a co-variate in the analytical models slightly altered the level of significance of the findings, but did not dramatically change the ranking of the top loci. One issue with this approach is that the algorithm was developed using 5mC profiles generated from fluorescence activated cell sorted (FACS) neurones and glia from cortical brain tissue, using the antibody NeuN. However, it has been reported in the literature that NeuN is not expressed by cerebellar Purkinje neurones (Guintivano et al. 2013). As such, this method may not be appropriate for determining neuron/glia proportions in cerebellar tissue, limiting its application to cross brain studies. Protocols for the separation of additional cellular subtypes from post-mortem brain tissue are now under development that will allow this analysis in the future (Guez-Barber et al. 2012, Krishnaswami et al. 2016, Mastroeni et al. 2017). These methods either use FACS with other antibodies to cell-specific antigens, or laser capture microdissection (LCM). However, one consideration for the future is that such discriminatory

analyses come at a price; for each additional cell type to be analysed by EWAS, this will significantly increase the cost of the experiment. Furthermore, FACS methodology is dependent on high quality tissue, robust antibodies suitable for identifying nuclei of target cells and currently has a low processing throughput, meaning this approach is not currently feasible in large cohorts.

6.3.2 Causality

Epigenetic changes in disease provide promise for future treatments due to their transient state. However the dynamic and temporal nature of epigenetic modifications means that it can be difficult to interpret epigenetic findings; changes in cytosine or histone modifications could be, in fact, a result of the disease process, rather than a cause. The issue of causality is especially challenging in the context of neurological disorders such as AD, where longitudinal EWAS analyses in the tissue of primary pathology (brain) are not possible. Therefore, addressing whether loci nominated from epigenome studies are causal in the disease process has been difficult, with two-step Mendelian randomisation providing the only potentially feasible method for establishing causality to date (Relton et al. 2012). However, recent developments in CRISPR-cas9 genome editing technology, have allowed researchers to specifically target the DNA methylation machinery making it applicable to manipulating epigenetic modifications (Sander et al. 2014), and will provide a superior method for causality determination in the future.

6.3.3 Cytosine modifications

One limitation of 5mC studies to date is the confounding effect of other cytosine modifications, such as 5hmC. 5hmC in particular has been found to be expressed at greater levels in the brain and in development (Kriaucionis et al. 2009, Wang et al. 2012), suggesting that this previously overlooked modification does have functional relevance. Due to the chemistry of traditional BS conversion methods, previous EWAS could not distinguish between 5mC and 5hmC. In Chapter 3 I addressed this issue by oxidising the DNA prior to BS conversion. The profiling of matched BS and OxBS treated samples simultaneously on the 450K array enabled the detection of both modifications throughout the genome. However, there are still nuances with this

methodology, for example the limited number of CpG sites profiled, and technical variation inherent in the profile resulting in some samples having a negative 5hmC level. The use of hMeDIP-seq would avoid this issue; however, hMeDIP-seq has a relatively low resolution and cannot quantitatively determine 5hmC abundance in a single-base resolution (Skvortsova et al. 2017). In addition, further studies are needed to interrogate whether other cytosine modifications present in the human genome (5fC and 5caC) are relevant in the study of AD.

6.3.4 *ANK1* coverage

Within this thesis, assays were designed to target key regions of pre-identified peaks in 5mC (Chapter 4) and histone modifications (Chapter 5) or assayed key CpG sites throughout the entire genome (Chapter 3). However, none of the methodologies employed in this thesis examine every cytosine or histone modification site in the *ANK1* gene. One potential solution is to use whole genome bisulfite sequencing (WGBS) to measure DNA modifications at single base resolution; however this would not distinguish between different DNA modifications. Third generation sequencing methodologies now exist that can detect different DNA modifications in native DNA, reducing the need for complex treatment strategies and allowing separate cytosine modification levels to be detected at every CpG site within the genome (Rhoads et al. 2015). However, these methodologies are very expensive, require large amounts of DNA and are analytically challenging due to large data volumes and new data types, meaning this form of epigenetic study is currently out of reach (Schadt et al. 2010). Similarly, looking to the future, the combination of ChIP and sequencing methodologies would allow the interrogation of H3K4me3 and H3K27me3 across the entire genome in AD; this method has been used successfully in AD for other histone modifications (Marzi et al. 2018).

6.3.5 *ANK1* expression

The *ANK1* gene is known to have multiple splice isoforms which are thought to be expressed in a tissue specific manner. Several studies have examined the expression levels of *ANK1* in human brain (De Jager et al. 2014, Lunnon et al. 2014, Mastroeni et al. 2017). Lunnon et al. (2014) used qPCR to assess the level of

expression of different *ANK1* isoforms and their relation to 5mC levels. Although 5mC has been traditionally described to be associated with gene repression, this is only really true in the context of CGIs within the 5' promoters of many constitutively expressed housekeeping control genes (Deaton et al. 2011). Recent evidence suggests that gene body methylation can be associated with gene activation or alternative gene splicing (Laurent et al. 2010), therefore it is possible that changes in 5mC at the *ANK1* locus could be driving the expression of specific transcript variants resulting in an altered protein isoform within the brains of AD patients. However, due to the sequence similarity in the transcript variants of *ANK1*, it is not possible to design specific qPCR assays to target individual variants, with assays available targeting multiple variant isoforms. Furthermore the level of *ANK1* expression in a normal brain is very low, no more than 12 transcripts per million (TPM) (Kapushesky et al. 2010), therefore huge sequencing depth would be needed to characterise the isoform expression profile in the brain. Long-read sequencing technology may hold the answer to this issue as it is able to perform long-read isoform sequencing of cDNA, which would allow the quantification of isoforms of any gene of interest (Rhoads et al. 2015).

6.4 Conclusions

These studies have now provided strong evidence for epigenetic dysfunction in AD, as well as identifying genes that could be used as new targets for pharmacological intervention. I have demonstrated epigenetic alterations in the *ANK1* gene in AD brain, which is also seen in some other, but not all dementias. However, many questions remain to be answered regarding the role the *ANK1* gene plays in AD. The most important of which are (1) whether AD-associated *ANK1* hypermethylation is a cause or a consequence of disease pathology, (2) whether the *ANK1* epigenetic changes are driven by a specific cellular population and (3) whether the epigenetic changes in *ANK1* in AD can be manipulated to create novel treatment strategies for the disease. Looking to the future, the latest advances in sequencing technology and cutting-edge genome manipulation tools will ultimately allow these questions to be answered.

APPENDIX 1 – ELUCIDATING NOVEL DISFUNCTIONAL PATHWAYS IN ALZHEIMER'S DISEASE BY INTERGRATING LOCI IDENTIFIED IN GENETIC AND EPIGENETIC STUDIES

Neuroepigenetics 6 (2016) 32–50

Contents lists available at ScienceDirect

Neuroepigenetics

journal homepage: www.elsevier.com/locate/nepig

Elucidating novel dysfunctional pathways in Alzheimer's disease by integrating loci identified in genetic and epigenetic studies

Adam R. Smith [a], Jonathan Mill [a,b], Rebecca G. Smith [a], Katie Lunnon [a,*]

[a] University of Exeter Medical School, RILD, Barrack Rd, University of Exeter, Devon, UK
[b] Institute of Psychiatry, Psychology and Neuroscience, King's College London, De Crespigny Park, London, UK

ARTICLE INFO

Article history:
Received 26 April 2016
Received in revised form 6 May 2016
Accepted 9 May 2016

Keywords:
Alzheimer's disease
AD
DNA methylation
GWAS
EWAS
Exome sequencing

ABSTRACT

Alzheimer's disease is a complex neurodegenerative disorder. A large number of genome-wide association studies have been performed, which have been supplemented more recently by the first epigenome-wide association studies, leading to the identification of a number of novel loci altered in disease. Twin studies have shown monozygotic twin discordance for Alzheimer's disease (Gatz et al., 2006), leading to the conclusion that a combination of genetic and epigenetic mechanisms is likely to be involved in disease etiology (Lunnon & Mill, 2013). This review focuses on identifying overlapping pathways between published genome-wide association studies and epigenome-wide association studies, highlighting dysfunctional synaptic, lipid metabolism, plasma membrane/cytoskeleton, mitochondrial, and immune cell activation pathways. Identifying common pathways altered in genetic and epigenetic studies will aid our understanding of disease mechanisms and identify potential novel targets for pharmacological intervention.

1. Introduction

Dementia encompasses a group of chronic neurodegenerative diseases that affected an estimated 44.4 million people worldwide in 2013. Because of an increasingly aging population, it is predicted that this figure will rise to an estimated 75.6 million by 2030 and 135.5 million by 2050 (Prince et al., 2013). The worldwide cost of dementia was estimated to be $604 billion in 2010 (Wimo & A.s.D. International, 2010). Alzheimer's disease (AD) is the most common form of dementia, accounting for ~60%–80% of cases worldwide (Lobo et al., 2000). AD is characterized by the accumulation of extracellular amyloid-β (Aβ) plaques, intracellular neurofibrillary tangles of hyperphosporylated tau, and widespread gliosis in the brain (Serrano-Pozo et al., 2011). Despite the progress that has been made in understanding the cellular pathology of AD, available treatments only temporarily alleviate some symptoms and do not modify the underlying disease process. By the time an individual becomes symptomatic, there are already considerable neuronal cell loss, plaque deposition, and tangle burden within the brain, which can appear up to 10 years before a clinical diagnosis is made (Jack et al., 2010). Reflecting the growing public health and socioeconomic burden of AD, there has been a year-on-year increase in the number of publications investigating the etiology of the disease (Fig. 1) as researchers seek novel disease-modifying treatments.

Although the neuropathology associated with AD has been well described, little is known about the mechanisms underlying disease onset and progression. Quantitative genetic analyses demonstrated high heritability estimates (58%–79%) for AD (Gatz et al., 2006), and thus, initial approaches to understanding etiology focused on uncovering a genetic contribution to disease susceptibility. In recent years, the recruitment of large cohorts and the relatively inexpensive cost of assessing genetic variation through genome-wide association studies (GWAS) have allowed the identification of multiple variants associated with an elevated risk of developing AD. Many of these genes have also been robustly associated with AD via subsequent meta-analyses (Harold et al., 2009; Hollingworth et al., 2011; Lambert et al., 2013a; Naj et al., 2011), and most recently, polygenic risk scores for AD have been developed (Escott-Price et al., 2015). Collectively, common single nucleotide polymorphisms (SNPs) are believed to only account for 33% of attributable risk (Ridge et al., 2013), and the mechanism behind their action remains largely unknown. Exome-sequencing projects have also identified other variants, for example, *TREM2* (Guerreiro et al., 2013), which have a larger effect size, yet these are relatively rare. In recent years, researchers have used epigenome-wide association studies (EWAS) to identify epigenetic changes in disease with the aim to elucidate additional mechanisms of pathology, which may provide a link to environmental factors.

* Corresponding author. Tel.: +44 1392 408 298.
E-mail address: k.lunnon@exeter.ac.uk (K. Lunnon).

http://dx.doi.org/10.1016/j.nepig.2016.05.001

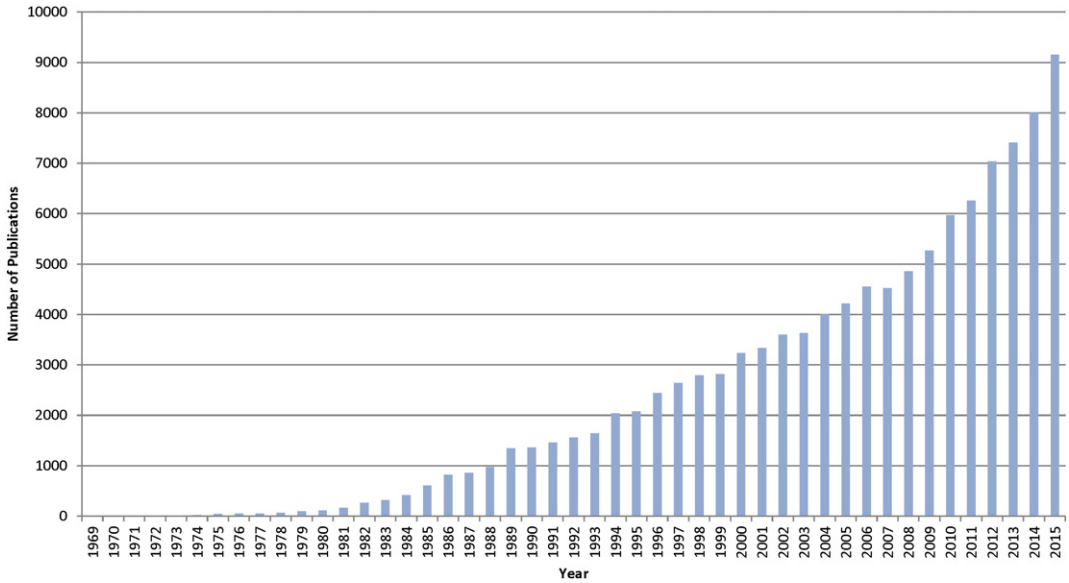


Fig. 1. The number of publications in the field of AD. Shown is the number of publications within the PubMed database that can be identified using the search term *Alzheimer's disease* between 1969 and 2015.

Epigenetic processes mediate the reversible regulation of gene expression, occurring independently of DNA sequence variation and acting principally through chemical modifications to DNA and nucleosomal histone proteins. Dynamic changes to the epigenome orchestrate a diverse range of important neurobiological and cognitive processes in the brain (Lunnon & Mill, 2013). DNA methylation is the best characterized and most stable epigenetic modification which modulates the transcription of mammalian genomes. This is due to its ability to be interrogated using archived genomic DNA resources, which are the focus of most human epidemiological epigenetic research to date (Lunnon & Mill, 2013). The methylation of a cytosine in a CpG dinucleotide by DNA methyltransferase enzymes forms 5-methylcytosine, which can disrupt the cell's transcriptional machinery by blocking the binding of transcription factors and attracting methyl-binding proteins that initiate chromatin compaction and bring about gene silencing (Klose & Bird, 2006). The predominant focus to date is methylation within CpG islands located within the 5′ promoters of many constitutively expressed housekeeping control genes. However, recent data suggest that the relationship between DNA methylation and transcription may be more complex, with gene body methylation and non-CpG methylation often being associated with active gene expression (Aran et al., 2011; Ball et al., 2009; Hellman & Chess, 2007; Lister et al., 2009) and alternative splicing (Flores et al., 2012; Lyko et al., 2010). The mechanisms involved in cytosine demethylation have also been studied; its demethylation by ten-eleven translocation (TET) enzymes leads to a stepwise change in the cytosine side chain state from methylated cytosine to hydroxymethylated cytosine (5-hmC), to formyl cytosine, to carboxyl cytosine, and finally back to unmodified cytosine by a yet unclassified enzyme/mechanism (Hill et al., 2014). Each of these intermediates may have their own effect on gene transcription, splicing, and subsequent protein function, and recent studies have shown 5-hmC to be at high levels in the brain (Nestor et al., 2012; Song et al., 2011), with variation across different anatomical regions (Lunnon et al., 2016). Recent advances in genomic technology have allowed the first genome-scale studies assessing methylomic variation (EWAS) in AD. These studies have identified AD-associated DNA methylomic variation at numerous loci in the cortex, with consistent findings across multiple independent study cohorts, in addition to brain-region–specific changes and blood DNA methylation signatures (Lunnon et al., 2014; De Jager et al., 2014). In addition, a recent paper by Yu et al. combined genetic and epigenetic findings by examining DNA methylation patterns across genes that have previously been nominated by GWAS, identifying several overlapping loci (Yu et al., 2014).

Although GWAS and EWAS analyses have identified multiple genes associated with AD, the extent to which common pathways are shared in the findings across studies has not yet been explored. This review aims to integrate the most robust findings from GWAS, exome sequencing studies, and EWAS performed to date in AD to highlight common molecular pathways, which could ultimately aid in the identification of novel pharmacological targets for the disease.

2. Methods

Using the publically available online search GWAS catalogue (https://www.ebi.ac.uk/gwas/search?query=Alzheimer%27s%20disease#association) and a P value cutoff of $P < 5 \times 10^{-8}$, we identified 45 unique GWAS in AD totaling 144 SNPs. We then removed studies based on poor sample size (<1000 total samples) as well as those studies that included samples that were non-European in origin. Following the study selection, SNPs in intronic regions were removed from the analysis. After filtering for associated disease outcome measures, including the terms *Dementia and core Alzheimer's disease neuropathological changes*, *Alzheimer's disease late onset*, *Alzheimer's disease*, *Psychosis and Alzheimer's disease*, *Alzheimer's disease age of onset*, *Alzheimer's disease biomarkers*, and *Neurofibrillary tangles*, we were left with 22 studies with 49 SNPs in 32 unique genes (Table 1A) (Harold et al., 2009; Hollingworth et al., 2011; Lambert et al., 2013a; Naj et al., 2011; Coon et al., 2007; Webster et al., 2008; Li et al., 2008; Kamboh et al., 2012a; Kamboh et al., 2012b; Lambert et al., 2009; Seshadri et al., 2010; Naj et al., 2010; Perez-Palma et al., 2014; Nelson et al., 2014; Abraham et al., 2008; Antunez et al., 2011; Hollingworth et al., 2012; Cruchaga et al., 2013; Hu et al., 2011; Lambert et al., 2013b; Beecham et al., 2014; Jonsson et al., 2013). Four genes were identified from exome sequencing studies (Guerreiro et al., 2013; Guerreiro et al., 2012; Cruchaga et al., 2014; Pottier et al., 2012) by performing a literature search in PubMed using the phrases *Alzheimer's disease* and *Exome sequencing* alone and in combination (Table 1B). Genes from EWAS were compiled from the 2014 publications by Lunnon et al. and De Jager et al. including probes with $P < 1 \times 10^{-7}$ (Lunnon et al., 2014; De Jager et al., 2014). The

Table 1A
Genes nominated from GWAS

Gene	SNPs identified	*P* value	GO annotation	Chromosome	Position	Title of article	First author and year	Reference
ABCA7	rs3764650	5.00E-17	GO:0,005,215 transporter activity GO:0,005,524 ATP binding GO:0,005,548 phospholipid transporter activity GO:0,016,887 contributes to ATPase activity	19	1,046,521	Common Variants at ABCA7, MS4A6A/ MS4A4E, EPHA1, CD33 and CD2AP Are Associated With Alzheimer's Disease	Hollingworth, 2011	(Hollingworth et al., 2011)
	rs4147929	1.00E-15		19	1,063,444	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
APOE, APOC1, TOMM40	rs4420638	1.00E-39	APOE GO:0,001,540 beta-amyloid binding IDA 11305869 GO:0,005,319 lipid transporter activity IDA 17305370 GO:0,005,515 protein binding IPI 12950167 GO:0,005,543 phospholipid binding	19	44,919,689	A High-Density Whole-Genome Association Study Reveals That APOE Is the Major Susceptibility Gene for Sporadic Late-Onset Alzheimer's Disease	Coon, 2007	(Coon et al., 2007)
		1.00E-39				Sorl1 as an Alzheimer's Disease Predisposition Gene?	Webster, 2008	(Webster et al., 2008)
		2.00E-44				Candidate Single-Nucleotide Polymorphisms From a Genomewide Association Study of Alzheimer Disease	Li, 2008	(Li et al., 2008)
		1.00E-12				Genome-Wide Association Analysis of Age-at-Onset in Alzheimer's Disease	Kamboh, 2012a	(Kamboh et al., 2012a)
		8.00E-149				Genome-Wide Association Study of Alzheimer's Disease	Kamboh, 2012b	(Kamboh et al., 2012b)
	rs2075650	2.00E-157	APOC1 GO:0,004,859 phospholipase inhibitor activity GO:0,005,504 fatty acid binding GO:0,031,210 phosphatidylcholine binding GO:0,055,102 lipase inhibitor activity GO:0,060,228 phosphatidylcholine-sterol *O*-acyltransferase activator activity	19	44,892,362	Genome-Wide Association Study Identifies Variants at CLU and PICALM Associated With Alzheimer's Disease	Harold, 2009	(Harold et al., 2009)
		2.00E-16				Genome-Wide Association Study Identifies Variants at CLU and CR1 Associated With Alzheimer's Disease	Lambert, 2009	(Lambert et al., 2009)
		1.00E-295				Genome-Wide Analysis of Genetic Loci Associated With Alzheimer Disease	Seshadri, 2010	(Seshadri et al., 2010)
		5.00E-36				Dementia Revealed: Novel Chromosome 6 Locus for Late-Onset Alzheimer Disease Provides Genetic Evidence for Folate-Pathway Abnormalities	Naj, 2010	(Naj et al., 2010)
		9.00E-116				Overrepresentation of Glutamate Signaling in Alzheimer's Disease: Network-Based Pathway Enrichment Using Meta-Analysis of Genome-Wide Association Studies	Perez-Palma, 2014	(Perez-Palma et al., 2014)

Table 1A (*continued*)

Gene	SNPs identified	*P* value	GO annotation	Chromosome	Position	Title of article	First author and year	Reference
		4.00E-13	TOMM40 GO:0,008,320 protein transmembrane transporter activity GO:0,015,288 porin activity			ABCC9 Gene Polymorphism Is Associated With Hippocampal Sclerosis of Aging Pathology	Nelson, 2014	(Nelson et al., 2014)
	rs6859	6.00E-14		19	44,878,777	A Genome-Wide Association Study for Late-Onset Alzheimer's Disease Using DNA Pooling	Abraham, 2008	(Abraham et al., 2008)
	rs157580	8.00E-89		19	44,892,009	The Membrane-Spanning 4-Domains, Subfamily A (MS4A) Gene Cluster Contains a Common Variant Associated With Alzheimer's Disease	Antunez, 2011	(Antunez et al., 2011)
	rs157582	9.00E-52		19	44,892,962	Genome-Wide Association Study of Alzheimer's Disease With Psychotic Symptoms	Hollingworth, 2012	(Hollingworth et al., 2012)
	rs769449	2.00E-18		19	44,906,745	GWAS of Cerebrospinal Fluid tau Levels Identifies Risk Variants for Alzheimer's Disease	Cruchaga, 2013	(Cruchaga et al., 2013)
BIN1	rs7561528	4.00E-14	GO:0,005,515 protein binding GO:0,032,403 protein complex binding GO:0,042,802 identical protein binding GO:0,046,982 protein heterodimerization activity GO:0,048,156 tau protein binding	2	127,132,061	Common Variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 Are Associated With Late-Onset Alzheimer's Disease	Naj, 2011	(Naj et al., 2011)
		6.00E-11				Genome-Wide Association Study of Alzheimer's Disease	Kamboh, 2012b	(Kamboh et al., 2012b)
	rs744373	3.00E-14		2	127,137,039	Common Variants at ABCA7, MS4A6A/ MS4A4E, EPHA1, CD33 and CD2AP Are Associated With Alzheimer's Disease	Hollingworth, 2011	(Hollingworth et al., 2011)
		1.00E-10				Meta-Analysis for Genome-Wide Association Study Identifies Multiple Variants at the BIN1 Locus Associated With Late-Onset Alzheimer's Disease	Hu, 2011	(Hu et al., 2011)
		2.00E-09				The Membrane-Spanning 4-Domains, Subfamily A (MS4A) Gene Cluster Contains a Common Variant Associated With Alzheimer's Disease	Antunez, 2011	(Antunez et al., 2011)
	rs6733839	7.00E-44		2	127,135,234	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
	rs12989701	3.00E-10		2	127,130,409	Meta-Analysis for Genome-Wide Association Study Identifies Multiple Variants at the BIN1 Locus Associated With Late-Onset Alzheimer's Disease	Hu, 2011	(Hu et al., 2011)
CASS4	rs7274581	3.00E-08	GO:0,000,155 phosphorelay sensor kinase activity GO:0,005,515 protein binding	20	56,443,204	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
CD2AP	rs10948363	5.00E-11	GO:0,005,200 structural constituent of cytoskeleton GO:0,005,515 protein binding	6	47,520,026	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci	Lambert, 2013a	(Lambert et al., 2013a)

(*continued on next page*)

Table 1A (*continued*)

Gene	SNPs identified	*P* value	GO annotation	Chromosome	Position	Title of article	First author and year	Reference
			GO:0,017,124 SH3 domain binding			for Alzheimer's Disease		
	rs9349407	9.00E-09		6	47,485,642	Common Variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 Are Associated With Late-Onset Alzheimer's Disease	Naj, 2011	(Naj et al., 2011)
CD33	rs3865444	2.00E-09	GO:0,004,872 receptor activity GO:0,005,515 protein binding GO:0,030,246 carbohydrate binding	19	51,224,706	Common Variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 Are Associated With Late-Onset Alzheimer's Disease	Naj, 2011	(Naj et al., 2011)
CELF1	rs10838725	1.00E-08	GO:0,000,166 nucleotide binding GO:0,000,900 translation repressor activity, nucleic acid binding GO:0,003,723 RNA binding IDA 16946708 GO:0,003,729 mRNA binding	11	47,536,319	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
CLU	rs9331896	3.00E-25	GO:0,005,515 protein binding GO:0,016,887 NOT ATPase activity GO:0,031,625 ubiquitin protein ligase binding GO:0,051,087 chaperone binding GO:0,051,787 misfolded protein binding	8	27,610,169	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
	rs2279590	6.00E-10		8	27,598,736	Genome-Wide Association Study Identifies Variants at CLU and CR1 Associated With Alzheimer's Disease	Lambert, 2009	(Lambert et al., 2009)
	rs11136000	9.00E-10		8	27,607,002	Genome-Wide Association Study Identifies Variants at CLU and PICALM Associated With Alzheimer's Disease	Harold, 2009	(Harold et al., 2009)
	rs569214	4.00E-08		8	27,630,273	The Membrane-Spanning 4-Domains, Subfamily A (MS4A) Gene Cluster Contains a Common Variant Associated With Alzheimer's Disease	Antunez, 2011	(Antunez et al., 2011)
CR1	rs6656401	6.00E-24	GO:0,001,851 complement component C3b binding GO:0,001,855 complement component C4b binding GO:0,001,861 complement component C4b receptor activity GO:0,004,877 complement component C3b receptor activity	1	207,518,704	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
	rs3818361	4.00E-14		1	207,611,623	Common Variants at ABCA7, MS4A6A/ MS4A4E, EPHA1, CD33 and CD2AP Are Associated With Alzheimer's Disease	Hollingworth, 2011	(Hollingworth et al., 2011)
	rs6656401	3.00E-10		1	207,518,704	Genome-Wide Association Study Identifies Variants at CLU and CR1 Associated With Alzheimer's Disease	Lambert, 2009	(Lambert et al., 2009)
	rs6701713	5.00E-10		1	207,612,944	Common Variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 Are Associated With Late-Onset Alzheimer's Disease	Naj, 2011	(Naj et al., 2011)
FERMT2	rs17125944	8.00E-09	GO:0,005,515 protein binding GO:0,005,547 phosphatidylinositol-	14	52,933,911	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)

Table 1A (*continued*)

Gene	SNPs identified	*P* value	GO annotation	Chromosome	Position	Title of article	First author and year	Reference
			3,4,5-trisphosphate binding					
FRMD4A	rs7081208	1.00E-10	GO:0,030,674 protein binding, bridging	10	13,949,865	Genome-Wide Haplotype Association Study Identifies the FRMD4A Gene as a Risk Locus for Alzheimer's Disease	Lambert, 2013b	(Lambert et al., 2013b)
GLIS3	rs514716	3.00E-09	GO:0,000,978 RNA polymerase II core promoter proximal region sequence-specific DNA binding sequence-specific DNA binding transcription factor activity involved in positive regulation of transcription GO:0,001,078 RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity involved in negative regulation of transcription	9	3,929,424	GWAS of Cerebrospinal Fluid tau Levels Identifies Risk Variants for Alzheimer's Disease	Cruchaga, 2013	(Cruchaga et al., 2013)
HLA-DRB5	rs9271192	3.00E-12	GO:0,042,605 peptide antigen binding	6	32,610,753	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
INPP5D	rs35349669	3.00E-08	GO:0,004,445 inositol-polyphosphate 5-phosphatase activity GO:0,005,515 protein binding GO:0,017,124 SH3 domain binding GO:0,034,594 phosphatidylinositol trisphosphate phosphatase activity GO:0,051,425 PTB domain binding	2	233,159,830	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
MEF2C	rs190982	3.00E-08	GO:0,000,977 RNA polymerase II regulatory region sequence-specific DNA binding GO:0,000,978 RNA polymerase II core promoter proximal region sequence-specific DNA binding GO:0,000,980 RNA polymerase II distal enhancer sequence-specific DNA binding GO:0,000,981 sequence-specific DNA binding RNA polymerase II transcription factor activity GO:0,000,983 RNA polymerase II core promoter sequence-specific DNA binding transcription factor activity	5	88,927,603	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
MS4A4A	rs4938933	8.00E-12	GO:0,016,021 integral component of membrane	11	60,266,956	Common Variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 Are Associated With Late-Onset Alzheimer's Disease	Naj, 2011	(Naj et al., 2011)
MS4A4E, MS4A6A	rs610932	2.00E-14	MS4A4E GO:0,016,021 integral component of membrane	11	60,171,834	Common Variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP Are Associated With Alzheimer's Disease	Hollingworth, 2011	(Hollingworth et al., 2011)
	rs983392	6.00E-16	MS4A6A GO:0,016,021 integral component of membrane	11	60,156,035	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)

(*continued on next page*)

Table 1A (*continued*)

Gene	SNPs identified	*P* value	GO annotation	Chromosome	Position	Title of article	First author and year	Reference
MTHFD1L	rs11754661	2.00E-10	GO:0,004,329 formate-tetrahydrofolate ligase activity GO:0,004,477 NOT methenyltetrahydrofolate cyclohydrolase activity GO:0,004,488 NOT methylenetetrahydrofolate dehydrogenase (NADP+) activity GO:0,005,524 ATP binding	6	150,885,942	Dementia Revealed: Novel Chromosome 6 Locus for Late-Onset Alzheimer Disease Provides Genetic Evidence for Folate-Pathway Abnormalities	Naj, 2010	(Naj et al., 2010)
NME8	rs2718058	5.00E-09	GO:0,004,550 nucleoside diphosphate kinase activity IBA GO:0,005,524 ATP binding	7	37,801,932	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
PICALM	rs10792832	9.00E-26	GO:0,005,515 protein binding GO:0,005,545 1-phosphatidylinositol binding GO:0,030,276 clathrin binding GO:0,032,050 clathrin heavy chain binding	11	86,156,833	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
	rs561655	7.00E-11		11	86,089,237	Common Variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 Are Associated With Late-Onset Alzheimer's Disease	Naj, 2011	(Naj et al., 2011)
	rs3851179	1.00E-09		11	86,157,598	Genome-Wide Association Study Identifies Variants at CLU and PICALM Associated With Alzheimer's Disease	Harold, 2009	(Harold et al., 2009)
	rs536841	3.00E-09		11	86,076,782	The Membrane-Spanning 4-Domains, Subfamily A (MS4A) Gene Cluster Contains a Common Variant Associated With Alzheimer's Disease	Antunez, 2011	(Antunez et al., 2011)
	rs17817600	2.00E-08		11	85,966,428	Genome-Wide Association Study of Alzheimer's Disease	Kamboh, 2012b	(Kamboh et al., 2012b)
PTK2B	rs28834970	7.00E-14	GO:0,004,683 calmodulin-dependent protein kinase activity GO:0,004,713 protein tyrosine kinase activity GO:0,004,715 non-membrane spanning protein tyrosine kinase activity GO:0,004,871 signal transducer activity	8	27,337,604	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
PVRL2	rs6857	2.00E-62	GO:0,001,618 virus receptor activity GO:0,005,515 protein binding GO:0,015,026 coreceptor activity GO:0,042,802 identical protein binding GO:0,042,803 protein homodimerization activity	19	44,888,997	Genome-Wide Association Meta-analysis of Neuropathologic Features of Alzheimer's Disease and Related Dementias	Beecham, 2014	(Beecham et al., 2014)
SLC24A4, RIN3	rs10498633	6.00E-09	SLC24A4 GO:0,008,273 calcium, potassium:sodium antiporter activity GO:0,015,293 symporter activity RIN3 GO:0,005,096 GTPase activator activity GO:0,005,515 protein binding GO:0,017,112 Rab guanyl-nucleotide exchange factor activity GO:0,017,137 Rab GTPase binding	14	92,460,608	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)
SORL1	rs11218343	1.00E-14	GO:0,001,540 beta-amyloid binding GO:0,004,888 transmembrane signaling receptor activity GO:0,005,515 protein binding GO:0,030,169 low-density lipoprotein particle binding	11	121,564,878	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)

Table 1A (*continued*)

Gene	SNPs identified	*P* value	GO annotation	Chromosome	Position	Title of article	First author and year	Reference
			GO:0,030,306 ADP-ribosylation factor binding					
NCR2	rs6922617	4.00E-08	GO:0,004,888 transmembrane signaling receptor activity	6	41,368,363	GWAS of Cerebrospinal Fluid tau Levels Identifies Risk Variants for Alzheimer's Disease	Cruchaga, 2013	(Cruchaga et al., 2013)
TREM2	rs75932628	2.00E-12	GO:0,001,530 lipopolysaccharide binding GO:0,004,872 receptor activity GO:0,005,515 protein binding GO:0,042,834 peptidoglycan binding GO:0,070,891 lipoteichoic acid binding	6	41,161,514	Variant of TREM2 Associated With the Risk of Alzheimer's Disease	Jonsson, 2013	(Jonsson et al., 2013)
UTS2D	rs9877502	5.00E-09	GO:0,001,664 G-protein coupled receptor binding GO:0,005,179 hormone activity	3	190,951,729	GWAS of Cerebrospinal Fluid tau Levels Identifies Risk Variants for Alzheimer's Disease	Cruchaga, 2013	(Cruchaga et al., 2013)
ZCWPW1	rs1476679	6.00E-10	GO:0,008,270 zinc ion binding	7	100,406,823	Meta-Analysis of 74,046 Individuals Identifies 11 New Susceptibility Loci for Alzheimer's Disease	Lambert, 2013a	(Lambert et al., 2013a)

Shown are the genes identified from GWAS and their respective SNPs, associated *P* value, GO annotation, chromosome and genomic position, and the relevant study. Only SNPs with $P < 5 \times 10^{-8}$ were included.

2012 publication by Bakulsk et al. was excluded from the analysis based on sample size (Bakulski et al., 2012). Gene names were checked against quoted genomic location using the UCSC Genome Browser; only genes containing a probe of interest were included. The resulting gene list contained 48 unique genes that met the criteria for inclusion for our study (Table 2). Gene annotation for all genes of interest were taken from the Gene Ontology (GO) Consortium database, where available, and supplemented with information from the Entrez gene database. Two genes overlapped between GWAS and exome sequencing studies (*TREM2*, *SORL1*), and 1 gene overlapped between GWAS and EWAS (*BIN1*), bringing the total number of genes across all analyses to 81.

3. Pathways

The 81 genes identified were compared in terms of their molecular/cellular function and grouped by pathways in which the identified genes operate. By taking significant loci across multiple

Table 1B
Genes nominated from exome sequencing studies

Gene	SNPs identified	GO annotation	Chromosome	Position	Title of article	First author and year	Reference
TREM2	rs2234255	GO:0,001,530 lipopolysaccharide binding GO:0,004,872 receptor activity GO:0,005,515 protein binding GO:0,042,834 peptidoglycan binding GO:0,070,891 lipoteichoic acid binding	6	41,127,543	TREM2 Variants in Alzheimer's Disease	Guerreiro, 2013	(Guerreiro et al., 2013)
	rs147564421		6	41,129,100			
	rs2234253		6	41,129,105			
	rs142232675		6	41,129,133			
	rs201258663		6	41,129,195			
	rs75932628		6	41,129,252			
	#N/A		6	41,129,279			
SORL1	#N/A	GO:0,001,540 beta-amyloid binding GO:0,004,888 transmembrane signaling receptor activity GO:0,005,515 protein binding GO:0,030,169 low-density lipoprotein particle binding GO:0,030,306 ADP-ribosylation factor binding	11	1.21E + 08	High Frequency of Potentially Pathogenic SORL1 Mutations in Autosomal Dominant Early-Onset Alzheimer Disease	Pottier, 2012	(Pottier et al., 2012)
	#N/A		11	1.21E + 08			
	#N/A		11	1.21E + 08			
NOTCH3	rs10408676	GO:0,005,509 calcium ion binding GO:0,005,515 protein binding GO:0,019,899 enzyme binding	19	15,290,007	Exome Sequencing Reveals an Unexpected Genetic Cause of Disease: NOTCH3 Mutation in a Turkish Family With Alzheimer's disease	Guerreiro, 2012	(Guerreiro et al., 2012)
PLD3	rs145999145	GO:0,004,630 phospholipase D activity GO:0,005,515 protein binding GO:0,070,290 N-acylphosphatidylethanolamine-specific phospholipase D activity	19	40,877,595	Rare Coding Variants in the Phospholipase D3 Gene Confer Risk for Alzheimer's Disease	Cruchaga, 2014	(Cruchaga et al., 2014)

Shown are the genes identified from exome sequencing studies and their respective SNPs, GO annotation, chromosome and genomic position, and the relevant study.

Table 2
Genes nominated from EWAS

Gene	Probes identified	*P* value	Functional summary, GO annotation	Chromosome	Position	Title of article	First author and year	Reference
ABR	cg25018458	1.89E-10	GO:0,005,089 Rho guanyl-nucleotide exchange factor activity GO:0,005,096 GTPase activator activity	17	980,014	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon, 2014	(Lunnon et al., 2014)
ALDH16A1	cg20618448	1.16E-08	GO:0,004,029 aldehyde dehydrogenase (NAD) activity	19	49,962,324	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
ANK1	cg05066959	7.13E-14	GO:0,005,198 structural molecule activity GO:0,005,200 structural constituent of cytoskeleton GO:0,005,515 protein binding GO:0,008,093 cytoskeletal adaptor activity GO:0,019,899 enzyme binding GO:0,030,507 spectrin binding GO:0,051,117 ATPase binding	8	41,519,308	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
		1.24E-09		8	41,519,308	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	[29
	cg11823178	7.83E-14		8	41,519,399	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
		3.42E-11		8	41,519,399	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon, 2014	(Lunnon et al., 2014)
	cg16140558	1.85E-08		8	41,514,039	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon, 2014	(Lunnon et al., 2014)
ASGR2	cg18659586	1.06E-09	GO:0,004,873 asialoglycoprotein receptor activity GO:0,005,515 protein binding GO:0,030,246 carbohydrate binding	17	7,017,474	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
ATG16L2	cg21806242	3.71E-10	Regulation of autophagy	11	72,532,891	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
BCAR3	cg02342148	1.60E-08	GO:0,005,085 guanyl-nucleotide exchange factor activity GO:0,005,515 protein binding	1	94,145,223	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
BIN1	cg22883290	3.73E-08	GO:0,005,515 protein binding GO:0,032,403 protein complex binding GO:0,042,802 identical protein binding GO:0,046,982 protein heterodimerization activity GO:0,048,156 tau protein binding	2	127,800,646	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
CDA	cg26407544	2.10E-10	GO:0,001,882 nucleoside binding GO:0,004,126 cytidine deaminase activity GO:0,005,515 protein binding GO:0,008,270 zinc ion binding	1	20,945,355	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
CDH23, C10orf54	cg23968456	1.09E-08	GO:0,005,509 calcium ion binding GO:0,005,515 protein binding	10	73,521,631	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon, 2014	(Lunnon et al., 2014)
		3.97E-10		10	73,521,631	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
CXXC5	cg07354506	9.26E-08	GO:0,004,871 signal transducer activity GO:0,005,515 protein binding GO:0,008,134 transcription factor binding GO:0,008,270 zinc ion binding	5	139,048,148	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon, 2014	(Lunnon et al., 2014)
FOXK1	cg07180538	4.95E-08	GO:0,000,977 RNA polymerase II regulatory region	7	4,786,899	Alzheimer's Disease: Early Alterations in Brain DNA	De Jager, 2014	(De Jager et al., 2014)

Table 2 (*continued*)

Gene	Probes identified	*P* value	Functional summary, GO annotation	Chromosome	Position	Title of article	First author and year	Reference
			sequence-specific DNA binding GO:0,000,981 sequence-specific DNA binding RNA polymerase II transcription factor activity GO:0,005,515 protein binding			Methylation at ANK1, BIN1, RHBDF2 and Other Loci		
	cg25594100	2.54E-11		7	4,786,943	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
GMDS	cg07714812	3.81E-08	GO:0,005,515 protein binding GO:0,008,446 GDP-mannose 4,6-dehydratase activity GO:0,070,401 NADP+ binding	6	1,635,611	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
GUCY2D	cg04157161	7.80E-08	GO:0,004,383 guanylate cyclase activity GO:0,004,672 protein kinase activity GO:0,004,872 receptor activity GO:0,005,524 ATP binding	17	7,906,847	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
HMHA1	cg02308560	3.06E-08	GO:0,005,096 GTPase activator activity GO:0,005,515 protein binding GO:0,046,872 metal ion binding	19	1,071,176	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
ITM2C	cg18346707	3.30E-08	GO:0,001,540 beta-amyloid binding GO:0,005,515 protein binding GO:0,005,524 ATP binding	2	231,732,249	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
ITPRIPL2	cg16733298	5.24E-08	membrane-associated protein, intracellular calcium signaling	16	19,127,132	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
KCNN4	cg22904711	1.08E-08	GO:0,005,515 protein binding GO:0,005,516 calmodulin binding GO:0,015,269 calcium-activated potassium channel activity GO:0,016,286 NOT small conductance calcium-activated potassium channel activity	19	44,278,628	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
KDM2B	cg11724984	4.76E-09	GO:0,003,677 DNA binding GO:0,005,515 protein binding GO:0,008,270 zinc ion binding GO:0,019,843 rRNA binding GO:0,032,452 histone demethylase activity	12	121,890,864	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
LNX1	cg12114584	5.81E-13	GO:0,004,842 ubiquitin-protein transferase activity GO:0,005,515 protein binding GO:0,008,270 zinc ion binding GO:0,016,874 ligase activity GO:0,030,165 PDZ domain binding	4	54,518,744	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
LYVE1	cg18343862	4.96E-08	GO:0,004,872 receptor activity GO:0,004,888 transmembrane signaling receptor activity GO:0,005,515 protein binding GO:0,005,540 hyaluronic acid binding	11	10,590,003	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
MAFB	cg13579486	3.25E-08	GO:0,003,700 sequence-specific DNA binding transcription factor activity GO:0,005,515 protein binding GO:0,008,134 transcription factor binding GO:0,043,565 sequence-specific DNA binding	20	39,314,091	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
MCF2L	cg07883124	6.31E-12	GO:0,005,089 Rho guanyl-nucleotide exchange factor activity GO:0,005,545 1-phosphatidylinositol binding	13	113,634,042	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
	cg09448088	6.43E-10		13	113,635,690	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
MYO10	cg06742628	1.58E-10	GO:0,005,515 protein binding GO:0,005,516 calmodulin binding	5	16,886,424	Alzheimer's Disease: Early Alterations in Brain DNA	De Jager, 2014	(De Jager et al., 2014)

(*continued on next page*)

Table 2 (*continued*)

Gene	Probes identified	*P* value	Functional summary, GO annotation	Chromosome	Position	Title of article	First author and year	Reference
			GO:0,005,524 ATP binding GO:0,005,547 phosphatidylinositol-3,4,5-trisphosphate binding			Methylation at ANK1, BIN1, RHBDF2 and Other Loci		
MYO1C	cg05417607	2.25E-08	GO:0,003,774 motor activity GO:0,003,779 actin binding GO:0,005,102 receptor binding GO:0,005,515 protein binding GO:0,005,516 calmodulin binding	17	1,373,605	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon, 2014	(Lunnon et al., 2014)
		2.52E-08		17	1,373,605	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
PASK	cg25488284	3.50E-08	GO:0,004,674 protein serine/threonine kinase activity GO:0,004,871 signal transducer activity GO:0,005,515 protein binding	2	242,048,127	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon, 2014	(Lunnon et al., 2014)
PCNT	cg00621289	6.48E-08	GO:0,005,515 protein binding GO:0,005,516 calmodulin binding	21	47,855,916	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
	cg04147621	1.39E-08		21	47,856,020	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon, 2014	(Lunnon et al., 2014)
	cg23449541	9.39E-08		21	47,855,893	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon, 2014	(Lunnon et al., 2014)
PLAT	cg17693222	2.14E-08	GO:0,004,252 serine-type endopeptidase activity GO:0,005,515 protein binding	8	42,033,472	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
PLXNC1	cg12877335	2.36E-11	GO:0,004,872 receptor activity GO:0,005,102 receptor binding GO:0,005,515 protein binding	12	94,539,319	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
PODXL	cg08737189	3.37E-08	GO:0,005,515 protein binding	7	131,223,417	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
	cg19140834	1.22E-09		7	131,217,668	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
PSMA1	cg27443779	1.18E-08	GO:0,003,723 RNA binding GO:0,004,298 threonine-type endopeptidase activity GO:0,005,515 protein binding	11	14,664,793	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
PSTPIP1	cg11652496	2.57E-09	GO:0,003,779 actin binding GO:0,005,515 protein binding GO:0,019,903 protein phosphatase binding	15	77,324,526	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
RBM33	cg13639901	1.54E-13	GO:0,000,166 nucleotide binding GO:0,044,822 poly(A) RNA binding	7	155,556,590	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
	cg14430943	2.71E-10		7	155,556,652	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
RHBDF2	cg05810363	9.42E-10	GO:0,004,252 serine-type endopeptidase activity GO:0,019,838 growth factor binding	17	74,475,270	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon, 2014	(Lunnon et al., 2014)
		3.68E-10		17	74,475,270	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
	cg12163800	2.66E-08		17	74,475,355	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
	cg13076843	3.81E-08		17	74,475,294	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon, 2014	(Lunnon et al., 2014)

Table 2 (*continued*)

Gene	Probes identified	*P* value	Functional summary, GO annotation	Chromosome	Position	Title of article	First author and year	Reference
		1.68E-09		17	74,475,294	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
SH3PXD2A	cg19007269	2.96E-09	GO:0,005,515 protein binding GO:0,035,091 phosphatidylinositol binding	10	105,420,501	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
SIX3	cg22385702	4.47E-10	GO:0,000,980 RNA polymerase II distal enhancer sequence-specific DNA binding GO:0,001,205 RNA polymerase II distal enhancer sequence-specific DNA binding transcription factor activity involved in positive regulation of transcription GO:0,001,222 transcription corepressor binding GO:0,003,700 sequence-specific DNA binding transcription factor activity	2	45,175,881	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
SLC15A4	cg06653632	9.77E-08	GO:0,005,290 L-histidine transmembrane transporter activity GO:0,015,293 symporter activity	12	129,281,444	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon K	(Lunnon et al., 2014)
SLC16A3	cg07012687	6.19E-09	GO:0,005,515 protein binding GO:0,008,028 monocarboxylic acid transmembrane transporter activity GO:0,015,129 lactate transmembrane transporter activity GO:0,015,293 symporter activity GO:0,015,355 secondary active monocarboxylate transmembrane transporter activity	17	80,195,180	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
SPG7	cg03169557	7.95E-09	GO:0,004,222 metalloendopeptidase activity GO:0,005,515 protein binding GO:0,005,524 ATP binding GO:0,008,233 peptidase activity GO:0,008,270 zinc ion binding	16	89,598,950	Methylomic Profiling Implicates Cortical Deregulation of ANK1 in Alzheimer's Disease	Lunnon K	(Lunnon et al., 2014)
		3.99E-10		16	89,598,950	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
STK32C	cg25917732	9.98E-08	GO:0,004,672 protein kinase activity GO:0,004,674 protein serine/ threonine kinase activity GO:0,004,713 protein tyrosine kinase activity null GO:0,005,524 ATP binding GO:0,016,772 transferase activity, transferring phosphorus-containing groups	10	134,038,395	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
TFAP2E	cg17474422	1.56E-08	GO:0,000,977 RNA polymerase II regulatory region sequence-specific DNA binding GO:0,000,981 sequence-specific DNA binding RNA polymerase II transcription factor activity GO:0,042,803 protein homodimerization activity	1	36,039,866	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
TMEM18	cg21644387	7.25E-08	GO:0,003,677 DNA binding	2	663,024	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
TPRG1	cg04252044	2.42E-10	GO:0,042,802 identical protein binding	3	188,664,747	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
	cg12307200	3.06E-17		3	188,664,632	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)

(*continued on next page*)

Table 2 (*continued*)

Gene	Probes identified	*P* value	Functional summary, GO annotation	Chromosome	Position	Title of article	First author and year	Reference
TTC22	cg15645660	4.05E-08	Chaperone activity	1	55,247,356	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
VRTN	cg21207436	2.24E-09	GO:0,004,803 transposase activity GO:0,043,565 sequence-specific DNA binding	14	74,815,316	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
VWF	cg27041424	4.47E-10	GO:0,001,948 glycoprotein binding GO:0,002,020 protease binding GO:0,005,178 integrin binding GO:0,005,515 protein binding GO:0,005,518 collagen binding	12	6,232,979	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
WDR81	cg19803550	1.04E-08	GO:0,016,772 transferase activity, transferring phosphorus-containing groups	17	1,637,391	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)
ZZEF1	cg06753513	3.87E-12	GO:0,005,509 calcium ion binding GO:0,008,270 zinc ion binding	17	3,977,385	Alzheimer's Disease: Early Alterations in Brain DNA Methylation at ANK1, BIN1, RHBDF2 and Other Loci	De Jager, 2014	(De Jager et al., 2014)

Shown are the genes identified from EWAS and their respective probes, associated *P* value, GO annotation, chromosome and genomic position, and the relevant study. Only probes with $P < 1 \times 10^{-7}$ were included.

study designs, we identified 5 common pathways altered at the genetic and/or epigenetic level in AD: plasma membrane and cytoskeletal processes, lipid homeostasis, synaptic signaling, immune cell processes, and mitochondrial processes (Fig. 2). The largest number of genes fell into the functional group plasma membrane and cytoskeletal processes (n = 14); however, this could be due to the fact that this is a proportionally larger pathway and is therefore more likely to contain an associated gene by chance. Of the pathways we have identified, many of them have considerable overlap; for example, lipid processes are intrinsically linked to the plasma membrane which is composed of phospholipids and a large percentage of cholesterol. To better understand the overlap between GWAS and EWAS nominated genes, we looked at the cellular localization of genes from each type of study (Fig. 3). The 2 largest localization groups (cellular membrane and nucleus) were consistent between methodologies. This would, to some degree, be expected

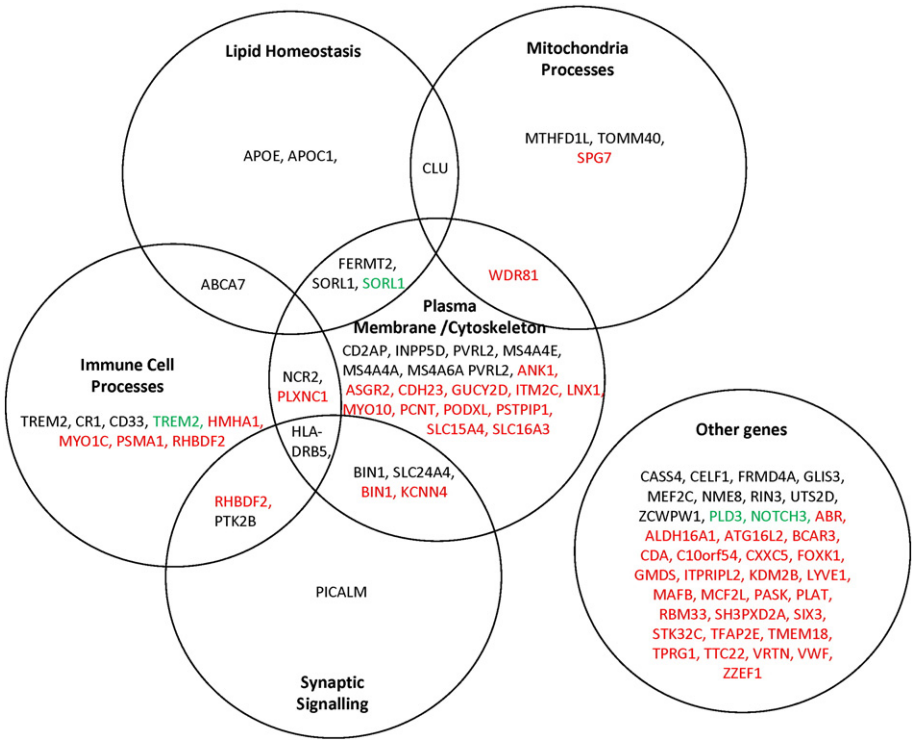


Fig. 2. Cellular pathways of genes identified from GWAS or EWAS. The 32 most significant GWAS loci identified (Table 1A) are shown in black. The 48 most significant EWAS loci (red) were identified from De Jager et al. (2014) and Lunnon et al. (2014) (Table 2). The 4 green loci were identified via exome sequencing from Cruchaga et al. (2014), Guerreiro et al. (2012), (2013), and Pottier et al. (2012) (Table 1B). We next compared the molecular and cellular pathways of the proteins encoded by these genes to look for functional overlap between studies.

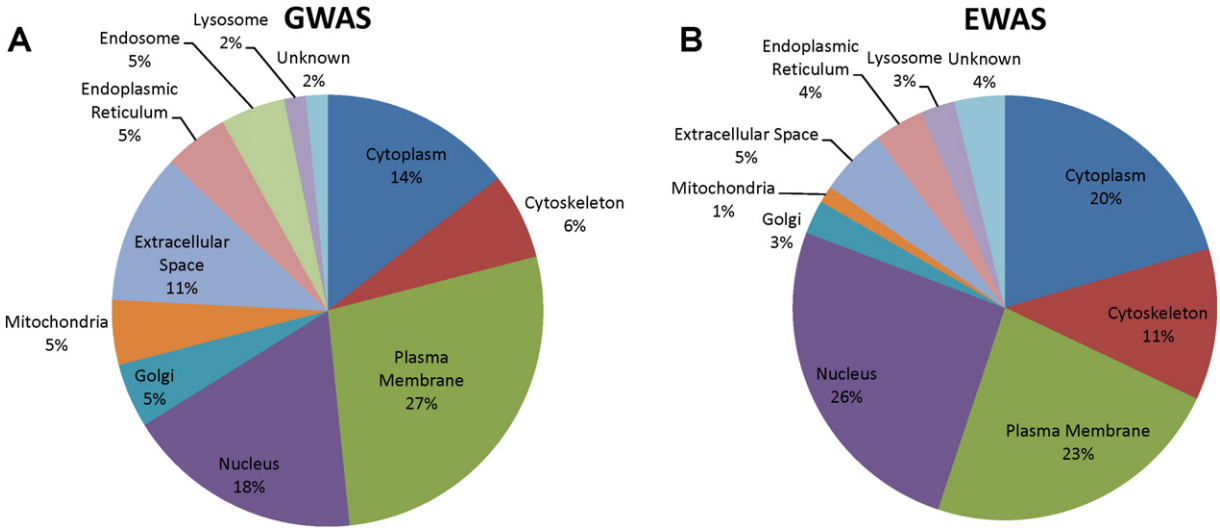


Fig. 3. Cellular locations of genes nominated from (A) GWAS and (B) EWAS.

because the majority of total proteins are involved in these locations and, in addition, current protein research has focused on these areas of the cell.

To provide a more structured approach to pathway analysis, all 81 genes were entered into the PANTHER pathway analysis using the enrichment analysis from gene ontology consortium (Mi et al., 2013). Fourteen biological process and 4 cellular component pathways were identified after passing Bonferroni correction. Most pathways reflected an interaction with Aβ or other AD pathology (Fig. 4). As the data for these genes were most likely collected from AD publications, the resulting pathways are not unexpected but are most likely to be limited.

3.1. Plasma membrane/cytoskeleton

This is the pathway which contained the largest number of associated genes from our analysis (n = 14). The plasma membrane insulates the intracellular components from the extracellular environment, as well as catalyzing the transport of specific compounds, including nutrients and ions. Phospholipids that make up the

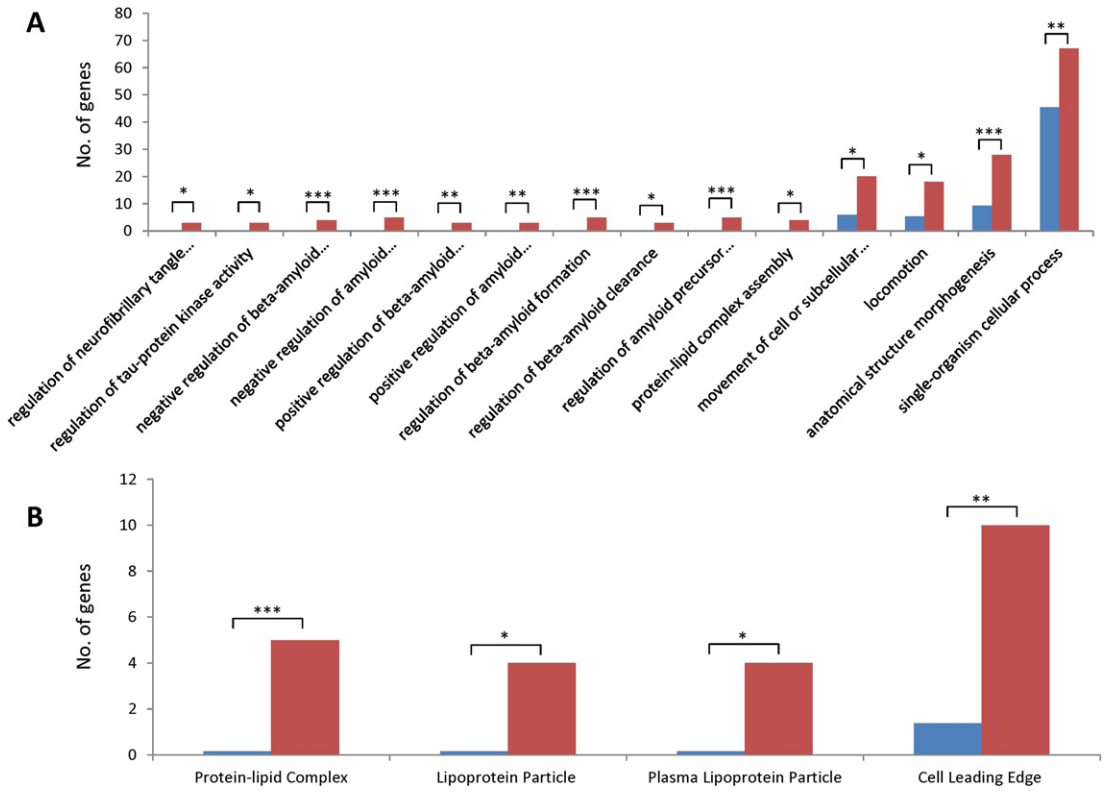


Fig. 4. Pathway enrichment analysis of 81 identified genes. (A) The Panther GO tool identified 14 significantly enriched biological processes pathways ($P < .05$) after Bonferroni correction. (B) Cellular component pathways identified by the Panther GO tool identified 4 cellular component pathways that were significantly enriched ($P < .05$) after Bonferroni correction. Blue columns represent expected number of genes in each pathway, whereas red represents the actual number of genes identified. Key: * = P<0.05, ** = P<0.01, *** = P<0.005.

membrane provide suitable fluidity and permeability. Alterations in the receptor function, membrane integrity, and membrane-dependent processes seen in AD have been reviewed by A. Farooqui et al. (Farooqui et al., 1995). The cytoskeleton provides contractility and couples biochemical responses with mechanical stresses in cells. It is vital in the movement of cellular machinery around the cell and to the membrane, as well as orchestrating the procedures needed for cellular movement and reshaping, a function specifically important to the microglial cells of the brain in the response to inflammation (Sheng et al., 1997). For an overview of cell mechanics and the cytoskeleton, see the review by Fletcher and Mullins 2010 (Fletcher & Mullins, 2010). The inability of neurons to regulate calcium homeostasis through cell surface ion channels is an aspect of AD pathogenesis that appears to be intimately involved in the dysfunction and death of neurons (Mattson, 2004). Familial AD mutations in *APP* and *PSEN1* support a role for perturbed calcium regulation in AD (Mattson, 2004). In addition, all of the enzymatic machinery responsible for the generation of the pathogenic Aβ plaque formation are plasma membrane based (Lukiw, 2013), suggesting that damage to the plasma membrane may be a key factor in the Aβ pathology typical of AD.

BIN1 has been nominated by both GWAS and EWAS, and in addition to its role in synaptic signaling, it also has a role in plasma membrane/cytoskeletal processes because it acts as an amphiphysin, which is known to promote caspase-independent apoptosis as well as play an important role in neuronal membrane organization (Wigge et al., 1997). Major learning defects and seizures have been linked to decreased expression of amphiphysins in murine brain (Di Paolo et al., 2002). In addition, altered expression of *BIN1* has been shown in aging mouse models of AD (Yang et al., 2008), providing further evidence for its role in AD pathology. Despite having no previous link to AD, *ANK1* is now the one of the strongest reported candidate genes in AD EWAS, with strong links to cell structure. *ANK1* was found to be hypermethylated in AD brain in 2 separate studies, including 1 with 2 independent validation cohorts (Lunnon et al., 2014; De Jager et al., 2014). The differentially methylated region in this gene spans at least 6 CpG sites and was significantly associated with neuropathology in cortical regions but not cerebellum or premortem blood (Lunnon et al., 2014), indicating tissue specificity of the differentially methylated region to regions of neuropathology. *ANK1* is found in multiple different isoforms, with some transcript variants specific to the brain (Gallagher et al., 1997) and some evidence for differential splicing in AD (Lunnon et al., 2014). As with *BIN1*, one of the main functions of *ANK1* is compartmentalization and maintenance of the plasma membrane, and it is possible that the altered expression of this gene could lead to neuronal membrane dysfunction in AD (Lunnon et al., 2014).

The *PVRL2* gene identified by GWAS encodes a single-pass type I membrane glycoprotein, which is one of the plasma membrane components of adherens junctions. Cell to cell connections brought about by adherens junctions are vital for effective neuronal signaling (Marambaud et al., 2002). Interestingly, Marambaud et al. used various immunological-based methods to investigate the *PSEN1*/γ-secretase system, where mutations are associated with familial AD, and showed that it disrupted adherens junctions in AD (Marambaud et al., 2002). Expression of *PVRL2* has been detected in many organs including the brain, and it was later suggested that it was associated with human longevity along with the AD GWAS nominated loci *TOMM40* and *APOE* (Lu et al., 2014). In addition, Elias-Sonnenschein et al. showed a significant correlation between the GWAS nominated locus *MS4A4A* and Aβ but not tau pathology in AD (Elias-Sonnenschein et al., 2013). Despite this, there is little to no research on the specific function of *MS4A4A*, although the gene product is associated with GO pathways that indicate that it is an integral component of the plasma membrane. Two other genes within the *MS4A* gene cluster have also been nominated via GWAS: *MS4A4E* and *MS4A6A* (Hollingworth et al., 2011; Lambert et al., 2013a). One recent study demonstrated that *MS4A6A* genotype and AD are associated with differential expression of isoform variants in blood and some brain regions (Proitsi et al., 2014).

3.2. Lipid homeostasis

Recent epidemiological, molecular, and biochemical evidence has strengthened the hypothesis that cholesterol is a risk factor for AD, and although cholesterol homeostasis in the brain is largely unexplored, new findings strongly support the involvement of cholesterol in both the generation and deposition of Aβ (Puglielli et al., 2003). Specifically, the quantity of cholesterol in the neuronal plasma membrane has been shown to make neurons more susceptible to the damage caused by Aβ in AD (Arispe & Doh, 2002). Other studies suggest that cholesterol acts directly on the amyloid cascade by promoting amyloidogenic processing of *APP* (Mattson, 2004). Interestingly, statins, which are a class of cholesterol-lowering drugs, decrease Aβ levels as well as plaque deposition in *APP* transgenic mouse models (Fassbender et al., 2001). In addition, high cholesterol levels and changes to cholesterol metabolism can increase the production of Aβ in cell culture and murine models (Puglielli et al., 2003). Three of the most significant genes from AD GWAS are associated with lipid metabolism (*APOE*, *APOC1*, *CLU*). *APOE* was first identified as a risk factor for AD in 1993 (Strittmatter et al., 1993) using immunostaining and genotyping analysis of 30 AD cases and 91 controls. Since 2006 and the wide application of GWAS to AD research (Grupe et al., 2007), the *APOE* polymorphism has been successfully replicated in several other studies (Coon et al., 2007; Abraham et al., 2008; Kramer et al., 2011; Logue et al., 2011; Meda et al., 2012; Ramanan et al., 2014), making *APOE* the most robust gene linked to late-onset AD (LOAD) risk to date. The proportion of genetic variance for LOAD risk attributed to *APOE* genotype is estimated to be 10%–20% (Slooter et al., 1998). *APOE* is a 299–amino acid glycoprotein and the major protein component of very low density lipoproteins, the major apolipoprotein in the brain (Puglielli et al., 2003), as well as having a functional role in cholesterol and triglyceride metabolism (Breslow et al., 1982). There are 3 *APOE* alleles that affect one's risk of AD (ε2, ε3, and ε4), in addition to age of onset (Roses & Allen, 1996). Of the 3 alleles, *APOE* ε2 demonstrates a protective effect, with an odds ratio of 0.3 for possessing one ε2 allele, whereas *APOE* ε4 is associated with a higher LOAD risk, with an odds ratio of 4.4 and 19.3, respectively for having 1 or 2 alleles (Corder et al., 1994), as well as a younger median age of dementia onset (Corder et al., 1994; Reiman et al., 2007). It has been suggested that the mutated *APOE* hinders clearance of soluble Aβ protein from the brain, leading to Aβ aggregation into fibrils. Furthermore, *APOE* has been shown to promote neurodegeneration by directing the toxic Aβ oligomers to synapses (Ramanan et al., 2014). However, a recent positron emission tomographic study to measure Aβ in 602 individuals found that the ε4 allele is neither necessary nor sufficient for the development of AD pathology (Ramanan et al., 2014).

SORL1 has been identified in several studies of AD using GWAS and exome sequencing methods; in addition, Yu et al. found epigenetic changes in this gene (Yu et al., 2014). It has many functional domains with different functions, including cargo transport, chaperone-like activity, signaling, and intracellular sorting (Jacobsen et al., 2001). When acting as a sorting receptor, the *SORL1* gene product protects *APP* from being directed to the endosome where it would be cleaved by β-secretase, producing Aβ (Louwersheimer et al., 2015). Furthermore, SORL1 can bind APOE, making *SORL1* an important component in the pathophysiology of AD (Elias-Sonnenschein et al., 2013).

3.3. Synaptic signaling

Synaptic dysfunction is possibly the best established of all the proposed pathological mechanisms for AD to date, as it shows clear

progression throughout the entire disease, including presymptomatic changes (Masliah et al., 2001). Early stages of AD are characterized by a 25%–35% decrease in numerical density of synapse per cortical region (Davies et al., 1987). There has also been evidence that the loss of synapses correlates with the soluble pool of cortical Aβ (Lue et al., 1999). Stereological and biochemical analyses have shown that the reduction in synaptic density within AD brain correlates with cognitive defects better than the traditional hallmarks of Aβ plaques and neurofibrillary tangles (Masliah et al., 2001).

We have identified 4 genes from GWAS and EWAS analyses of AD that have been linked to synaptic function. Two of these, *BIN1* and *PICALM*, have functions in vesicular trafficking. Specifically, studies have shown that the *BIN1* gene has roles in a number of specific pathways, including clathrin-mediated endocytosis which is an essential step in the intracellular trafficking of proteins and lipids such as nutrients, growth factors, and neurotransmitters in synapses (Cousin & Robinson, 2001; Dreyling et al., 1996; Wigge & McMahon, 1998). Originally identified as a tumor suppressor (Sakamuro et al., 1996), the *BIN1* gene product is expressed most abundantly in brain and muscle (Wechsler-Reya et al., 1997), with several alternatively spliced brain-specific isoforms. *BIN1* is one of the few genes reproducibly identified by GWAS that does not fall near or within the *APOE* locus; in addition, it is the only gene in our analysis to be significantly associated with AD in both GWAS and EWAS.

Like *BIN1*, *PICALM* is also involved in clathrin-mediated endocytosis (Dreyling et al., 1996). PICALM directs the trafficking of the VAMP2 protein. VAMP2 is a SNARE protein that plays a key role in the fusion of vesicles to the presynaptic membrane, allowing neurotransmitter release into the synapse, a process essential to neuronal function (Harel et al., 2008). *PICALM* has been robustly identified as a risk factor for AD via GWAS (Harold et al., 2009; Seshadri et al., 2010); however, AD-linked SNPs identified in *PICALM* may still be affected by *APOE* genotype because of the large amount of attenuation seen when adjusted for *APOE* status (Hu et al., 2011). Jun et al. have also reported this interaction observing that genotypes of *PICALM* conferred risk predominantly in *APOE* ε4–positive participants, providing strong evidence for a synergistic effect (Jun et al., 2010). PICALM is also thought to affect amyloid precursor protein processing via endocytic pathways (Harold et al., 2009).

As a previously known risk factor gene for AD (Lambert et al., 2013a), *PTK2B* was shown via network analyses to be linked to *RHBDF2*, *ANK1*, and *RPL13*, which were recently nominated from EWAS, providing further evidence for a role in AD pathology (De Jager et al., 2014). *PTK2B* has a number of roles including the induction of long-term potentiation of nerve cells, a central process of memory formation; cell migration; and synaptic function (Lambert et al., 2013a).

3.4. Immune cell dysfunction (astrocytes, oligodendrocytes, and microglia)

There is a widely accepted link between inflammation, the immune system, and AD pathology (Adriana Martorana et al., 2012; Akiyama et al., 2000; Monson et al., 2014; Salminen et al., 2009; Tuppo & Arias, 2005); more specifically, the inflammation seen in AD has been proposed to exacerbate symptoms (Akiyama et al., 2000). Microglia, which are the brain's resident macrophages, have been shown to increase their viability by 22.0%–29.4% in response to fibrillar Aβ deposits of 0.2 to 5.0 μmol/L, which are commonly seen in AD. Oligomeric Aβ at a dose of 5.0 μmol/L results in cytotoxic microglia (Pan et al., 2011) and ultimately leads to synaptic degeneration and neuronal death (Barger & Basile, 2001). However, relatively few genes that have shown robust associations with AD have been directly linked with inflammation or immune functions. Most noteworthy, a rare variant in *TREM2* was recently recognized by a number of AD exome sequencing studies and GWAS (Guerreiro et al., 2013; Jonsson et al., 2013; Forabosco et al., 2013; Neumann & Daly, 2012). *TREM2* encodes an innate immune system receptor on the surface of microglial cells within the brain. With the signaling counterpart DAP12 (also called TYROBP), TREM2 forms a molecular complex that promotes phagocytosis of bacteria (N'Diaye et al., 2009). Work by Takahashi et al. has shown that TREM2 also has a role in the clearance of apoptotic neurones due to its ability to increase migration and phagocytosis of microglia (Takahashi et al., 2005). Recently, 1 study demonstrated correlation in *TREM2* and *CD33* gene expression in AD (Chan et al., 2015). As *CD33* has also been nominated in various AD GWAS (Hollingworth et al., 2011; Naj et al., 2011; Kamboh et al., 2012b), this provides further evidence for an overlap of AD gene pathways in disease. As described above, recent protein-protein interaction data also demonstrated that several EWAS nominated loci (*ANK1*, *RHBDF2*, *PICLAM*) have a functional link to *PTK2B* (De Jager et al., 2014). *PTK2B* is an AD risk factor gene that plays a key role in the signaling cascade involved in the modulation of microglial and infiltrating macrophage cell activation (De Jager et al., 2014).

A further gene related to immune function is *RHBDF2*, identified by EWAS. Differentially methylated CpG sites close to the *RHBDF2* gene were identified in 2 independent EWAS (Lunnon et al., 2014; De Jager et al., 2014), with recent studies showing that this increases *RHBDF2* expression in AD brain (De Jager et al., 2014). *RHBDF2* transports TNFα converting enzyme (TACE, also called ADAM17), which is necessary for the release of TNFα from the cell surface (Adrain et al., 2012). *RHBDF2* absence in mice affects the release of TNFα from the cell surface (Siggs et al., 2012) and therefore impairs systemic immune responses to pathogens (McIlwain et al., 2012), although the brain phenotype has yet to be researched.

3.5. Mitochondrial processes

Mitochondrial dysfunction is one of the most prominent characteristics of AD in both the brain and the periphery (Devall et al., 2016; Devall et al., 2014; Lunnon et al., 2012), with *TOMM40*, one of the most robust genes identified from GWAS, associated with mitochondrial function. This gene is located approximately 2 kilobases downstream from *APOE*, and because of the locality of these 2 genes, there is strong linkage disequilibrium (LD) for *TOMM40* with the *APOE* locus (Feulner et al., 2010); hence, many studies have failed to find an association of *TOMM40* in AD after adjusting for *APOE* genotype (Ramanan et al., 2014; Wijsman et al., 2011; Yu et al., 2007). However, 1 study reports *TOMM40* as a possible risk factor of AD independent of *APOE* (Lutz et al., 2010). Specifically, this study found a poly-T track mutation in *TOMM40* that acts independently of *APOE* genotype, which has also seen been reported in another independent study (Cruchaga et al., 2011). In addition to increasing risk of developing AD, *TOMM40* has also been linked to an earlier age of onset for the disease (Roses et al., 2009). Other studies also suggest that *TOMM40* provides an additional risk for AD, in addition to *APOE* (Potkin et al., 2009; Takei et al., 2009). However, until the extent of the LD between *TOMM40* and *APOE* is fully characterized, it will be difficult to pinpoint the exact effect the *TOMM40* mutation has on LOAD pathogenesis.

CLU has various nuclear and mitochondrial isoforms and is thought to regulate the rate of cell proliferation. *CLU* has been consistently replicated across many GWAS and holds a strong association with AD (Harold et al., 2009; Lambert et al., 2009; Jun et al., 2010; Wijsman et al., 2011). The nuclear isoforms result in the promotion of apoptosis, whereas mitochondrial isoforms of *CLU* suppress BAX-dependent release of cytochrome c into the cytoplasm and inhibit apoptosis (Zhang et al., 2005). As an increased level of apoptosis in the brain is seen in AD, it could suggest a role of *CLU* mutations in pathogenesis (Behl, 2000). *SPG7* was identified by EWAS and encodes a mitochondrial metalloprotease protein. Mitochondrial proteases degrade misfolded and nonassembled polypeptides. They also regulate the activity of specific substrates by mediating essential processing steps. These proteases have been hypothesized to play a role in

neurodegenerative diseases by affecting neuronal maintenance and axonal function (Martinelli & Rugarli, 2010).

4. Discussion

The use of GWAS to identify common disease variants in AD has been at the forefront of research to understanding disease etiology for 10 years. More recently, the falling cost of exome and whole genome sequencing has identified rarer variants with a larger effect size. However, only 3 EWAS have been reported in AD to date (Lunnon et al., 2014; De Jager et al., 2014; Bakulski et al., 2012), which have solely focused on DNA methylation, although further studies are highly anticipated. Of all the genes identified from GWAS and EWAS in AD, only 1 locus was found to be overlapping between these 2 methodologies (*BIN1*).

As with any pathway identification analysis, there are caveats to our method. Some pathways are significantly larger than others containing more genes; therefore, using this method, we are more likely to find associated genes in these pathways over others. Secondly, cellular pathways that contain a gene which is either genetically or epigenetically altered may still be able to function normally, as similar proteins could "step-in" to fulfill the lost functionality. Thirdly, in our analysis, we did not filter our results based on loss of function SNPs or reduced expression; therefore, despite the alterations in AD, the genes we have identified may well have no change in their functionality. Fourthly, AD is characterized by neuronal cell loss and gliosis, and thus, the findings from EWAS may simply represent an alteration in cellular abundance, and although EWAS can apply cell-specific corrections to methylation data (Guintivano et al., 2013), this was not included in our analysis. The ability to look at single-cell epigenetic profiles in disease would allow researchers to conclusively quantify changes that occur at both cellular and disease-state levels; however, single-cell isolation in postmortem tissue, via laser capture microdissection or florescent-assisted cell sorting, currently represents a considerable challenge to the field. Finally, epigenetic research in AD is still in its early stages with only 2 EWAS included in our analysis; this coupled with the fact that current methylation data are the sum of 2 different cytosine modifications (5-methylcytosine and 5-hmC) means that we may have an underrepresentation of significant EWAS genes in AD. A further caveat of epigenetic studies compared with genetic studies is that causality is more difficult to establish, and thus, further studies examining the functional role of nominated EWAS loci are warranted.

5. Conclusion

Looking at the most significant genetic and epigenetic findings in AD to date, we have identified several pathways that require further exploration and could ultimately aid in our understanding of AD etiology. Well-characterized clinical cohorts will also allow the identification of further rare variants of AD, whereas advances in methodologies are also allowing the identification of other epigenetic marks, such as histone modifications and other DNA modifications at single-nucleotide resolution (Lunnon & Mill, 2013). A number of recent studies have demonstrated altered global levels of 5-hmC in AD brain (Condliffe et al., 2014; Chouliaras et al., 2013); however, studies to investigate loci-specific 5-hmC changes in AD are yet to be published. There is also the potential for further disease mechanisms to be identified from current studies as research moves to integrate GWAS and EWAS data in the same data sets to identify *cis* methylation quantitative trait loci. Ultimately, integrating genomic and epigenomic data with other "omic" modalities will allow the identification of novel dysfunctional pathways in disease (Lunnon & Mill, 2013).

Acknowledgements

This work was funded by a grant from Bristol Research into Alzheimer's and Care of the Elderly and the Alzheimer's Society (grant AS-PG-14-038) to KL.

References

Abraham, R., et al., 2008. A genome-wide association study for late-onset Alzheimer's disease using DNA pooling. BMC Med. Genet. 1, 44.
Adrain, C., et al., 2012. Tumor necrosis factor signaling requires iRhom2 to promote trafficking and activation of TACE. Science 335 (6065), 225–228.
Adriana Martorana, M.B., Buffa, S., Pellicanò, M., Caruso, C., Colonna-Romano, G.C.a.G., 2012. Immunosenescence, inflammation and Alzheimer's disease. Longev. Lifespan 1.
Akiyama, H., Barnum, S., Bradt, B., Bauer, J., Cole, G.M., Cooper, N.R., Eikelenboom, P., Emmerling, M., Fiebich, B.L., Finch, C.E., Frautschy, S., Griffin, W.S.T., Hampel, H., Hull, M., Landreth, G., Lue, L.–.F., Mrak, R., Mackenzie, I.R., McGeer, P.L., O'Banion, M.K., Pachter, J., Pasinetti, G., Plata–Salaman, C., Rogers, J., Rydel, R., Shen, Y., Streit, W., Strohmeyer, R., Tooyoma, I., Van Muiswinkel, F.L., Veerhuis, R., Walker, D., Webster, S., Wegrzyniak, B., Wenk, G., Wyss–Coray, T., 2000. Inflammation and Alzheimer's disease. Neurobiol. Aging 21, 383–421.
Antunez, C., et al., 2011. The membrane-spanning 4-domains, subfamily a (MS4A) gene cluster contains a common variant associated with Alzheimer's disease. Genitourin. Med. 3 (5), 33.
Aran, D., et al., 2011. Replication timing-related and gene body-specific methylation of active human genes. Hum. Mol. Genet. 20 (4), 670–680.
Arispe, N., Doh, M., 2002. Plasma membrane cholesterol controls the cytotoxicity of Alzheimer's disease AβP (1-40) and (1-42) peptides. FASEB J. 16 (12), 1526–1536.
Bakulski, K.M., et al., 2012. Genome-wide DNA methylation differences between late-onset Alzheimer's disease and cognitively normal controls in human frontal cortex. J. Alzheimers Dis. 29 (3), 571–588.
Ball, M.P., et al., 2009. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat. Biotechnol. 27 (4), 361–368.
Barger, S.W., Basile, A.S., 2001. Activation of microglia by secreted amyloid precursor protein evokes release of glutamate by cystine exchange and attenuates synaptic function. J. Neurochem. 76 (3), 846–854.
Beecham, G.W., et al., 2014. Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. PLoS Genet. 10 (9), e1004606.
Behl, C., 2000. Apoptosis and Alzheimer's disease. J. Neural Transm. 107 (11), 1325–1344.
Breslow, J.L., et al., 1982. Studies of familial type III hyperlipoproteinemia using as a genetic marker the apoE phenotype E2/2. J. Lipid Res. 23 (8), 1224–1235.
Chan, G., et al., 2015. CD33 modulates TREM2: convergence of Alzheimer loci. Nat. Neurosci. 18 (11), 1556–1558.
Chouliaras, L., et al., 2013. Consistent decrease in global DNA methylation and hydroxymethylation in the hippocampus of Alzheimer's disease patients. Neurobiol. Aging 34 (9), 2091–2099.
Condliffe, D., et al., 2014. Cross-region reduction in 5-hydroxymethylcytosine in Alzheimer's disease brain. Neurobiol. Aging 35 (8), 1850–1854.
Coon, K.D., P.D.A.J.M., Craig, D.W., et al., 2007. A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. J. Clin. Psychol. 68 (4), 613–618.
Corder, E.H., et al., 1994. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. Nat. Genet. 7 (2), 180–184.
Cousin, M.A., Robinson, P.J., 2001. The dephosphins: dephosphorylation by calcineurin triggers synaptic vesicle endocytosis. Trends Neurosci. 24 (11), 659–665.
Cruchaga, C., et al., 2011. Association and expression analyses with single-nucleotide polymorphisms in TOMM40 in Alzheimer disease. Arch. Neurol. 68 (8), 1013–1019.
Cruchaga, C., et al., 2013. GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease. Neuron 78 (2), 256–268.
Cruchaga, C., et al., 2014. Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease. Nature 505 (7484), 550–554.
Davies, C.A., et al., 1987. A quantitative morphometric analysis of the neuronal and synaptic content of the frontal and temporal cortex in patients with Alzheimer's disease. J. Neurol. Sci. 78 (2), 151–164.
De Jager, P.L., et al., 2014. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. Nat. Neurosci. 17 (9), 1156–1163.
Devall, M., Mill, J., Lunnon, K., 2014. The mitochondrial epigenome: a role in Alzheimer's disease? Epigenomics 6 (6), 665–675.
Devall, M., et al., 2016. Epigenetic regulation of mitochondrial function in neurodegenerative disease: new insights from advances in genomic technologies. Neurosci. Lett. Feb 10. pii: S0304-3940(16)30081-7.
Di Paolo, G., et al., 2002. Decreased synaptic vesicle recycling efficiency and cognitive deficits in amphiphysin 1 knockout mice. Neuron 33 (5), 789–804.
Dreyling, M.H., et al., 1996. The t(10;11)(p13;q14) in the U937 cell line results in the fusion of the AF10 gene and CALM, encoding a new member of the AP-3 clathrin assembly protein family. Proc. Natl. Acad. Sci. U. S. A. 93 (10), 4804–4809.
Elias-Sonnenschein, L.S., et al., 2013. Genetic loci associated with Alzheimer's disease and cerebrospinal fluid biomarkers in a Finnish case-control cohort. PLoS One 8 (4), e59676.
Escott-Price, V., et al., 2015. Common polygenic variation enhances risk prediction for Alzheimer's disease. Brain 138 (Pt 12), 3673–3684.

Farooqui, A., Wells, K., Horrocks, L., 1995. Breakdown of membrane phospholipids in Alzheimer disease. Mol. Chem. Neuropathol. 25 (2–3), 155–173.
Fassbender, K., et al., 2001. Simvastatin strongly reduces levels of Alzheimer's disease β-amyloid peptides Aβ42 and Aβ40 in vitro and in vivo. Proc. Natl. Acad. Sci. 98 (10), 5856–5861.
Feulner, T.M., et al., 2010. Examination of the current top candidate genes for AD in a genome-wide association study. Mol. Psychiatry 15 (7), 756–766.
Fletcher, D.A., Mullins, R.D., 2010. Cell mechanics and the cytoskeleton. Nature 463 (7280), 485–492.
Flores, K., et al., 2012. Genome-wide association between DNA methylation and alternative splicing in an invertebrate. BMC Genomics 13, 480.
Forabosco, P., et al., 2013. Insights into TREM2 biology by network analysis of human brain gene expression data. Neurobiol. Aging 34 (12), 2699–2714.
Gallagher, P.G., et al., 1997. Structure and organization of the human ankyrin-1 gene: basis for complexity of pre-mRNA processing. J. Biol. Chem. 272 (31), 19220–19228.
Gatz, M., et al., 2006. Role of genes and environments for explaining Alzheimer disease. Arch. Gen. Psychiatry 63 (2), 168–174.
Grupe, A., et al., 2007. Evidence for novel susceptibility genes for late-onset Alzheimer's disease from a genome-wide association study of putative functional variants. Hum. Mol. Genet. 16 (8), 865–873.
Guerreiro, R.J., et al., 2012. Exome sequencing reveals an unexpected genetic cause of disease: NOTCH3 mutation in a Turkish family with Alzheimer's disease. Neurobiol. Aging 33 (5) (1008e17-23).
Guerreiro, R., et al., 2013. TREM2 variants in Alzheimer's disease. N. Engl. J. Med. 368 (2), 117–127.
Guintivano, J., Aryee, M., Kaminsky, Z., 2013. A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. Epilepsia 8 (3), 290–302.
Harel, A., et al., 2008. Evidence for CALM in directing VAMP2 trafficking. Traffic 9 (3), 417–429.
Harold, D., et al., 2009. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat. Genet. 41 (10), 1088–1093.
Hellman, A., Chess, A., 2007. Gene body-specific methylation on the active X chromosome. Science 315 (5815), 1141–1143.
Hill, P.W.S., Amouroux, R., Hajkova, P., 2014. DNA demethylation, Tet proteins and 5-hydroxymethylcytosine in epigenetic reprogramming: an emerging complex story. Genomics 104 (5), 324–333.
Hollingworth, P., et al., 2011. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nat. Genet. 43 (5), 429–435.
Hollingworth, P., et al., 2012. Genome-wide association study of Alzheimer's disease with psychotic symptoms. Mol. Psychiatry 17 (12), 1316–1327.
Hu, X., et al., 2011. Meta-analysis for genome-wide association study identifies multiple variants at the BIN1 locus associated with late-onset Alzheimer's disease. PLoS One 6 (2), e16616.
Jack Jr., C.R., et al., 2010. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. Lancet Neurol. 9 (1), 119–128.
Jacobsen, L., et al., 2001. Activation and functional characterization of the mosaic receptor SorLA/LR11. J. Biol. Chem. 276 (25), 22788–22796.
Jonsson, T., et al., 2013. Variant of TREM2 associated with the risk of Alzheimer's disease. N. Engl. J. Med. 368 (2), 107–116.
Jun, G., et al., 2010. Meta-analysis confirms CR1, CLU, and PICALM as alzheimer disease risk loci and reveals interactions with APOE genotypes. Arch. Neurol. 67 (12), 1473–1484.
Kamboh, M.I., et al., 2012a. Genome-wide association analysis of age-at-onset in Alzheimer's disease. Mol. Psychiatry 17 (12), 1340–1346.
Kamboh, M.I., et al., 2012b. Genome-wide association study of Alzheimer's disease. Transl. Psychiatry 2, e117.
Klose, R.J., Bird, A.P., 2006. Genomic DNA methylation: the mark and its mediators. Trends Biochem. Sci. 31 (2), 89–97.
Kramer, P.L., et al., 2011. Alzheimer disease pathology in cognitively healthy elderly: a genome-wide study. Neurobiol. Aging 32 (12), 2113–2122.
Lambert, J.C., et al., 2009. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat. Genet. 41 (10), 1094–1099.
Lambert, J.-C., et al., 2013a. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat. Genet. 45 (12), 1452–1458.
Lambert, J.C., et al., 2013b. Genome-wide haplotype association study identifies the FRMD4A gene as a risk locus for Alzheimer's disease. Mol. Psychiatry 18 (4), 461–470.
Li, H., et al., 2008. Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease. Arch. Neurol. 65 (1), 45–53.
Lister, R., et al., 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462 (7271), 315–322.
Lobo, A., et al., 2000. Prevalence of dementia and major subtypes in Europe: a collaborative study of population-based cohorts. Neurologic diseases in the elderly research group. Neurology 54 (11 Suppl 5), S4–S9.
Logue, M.W., et al., 2011. A comprehensive genetic association study of Alzheimer disease in African Americans. Arch. Neurol. 68 (12), 1569–1579.
Louwersheimer, E., et al., 2015. The influence of genetic variants in SORL1 gene on the manifestation of Alzheimer's disease. Neurobiol. Aging 36 (3), 1605.e13–1605.e20.
Lu, F., et al., 2014. Genetic variants in PVRL2-TOMM40-APOE region are associated with human longevity in a Han Chinese population. PLoS One 9 (6), e99580.
Lue, L.-F., et al., 1999. Soluble amyloid β peptide concentration as a predictor of synaptic change in Alzheimer's disease. Am. J. Pathol. 155 (3), 853–862.
Lukiw, W.J., 2013. Alzheimer's disease (AD) as a disorder of the plasma membrane. Front. Physiol. 4, 24.
Lunnon, K., Mill, J., 2013. Epigenetic studies in Alzheimer's disease: current findings, caveats, and considerations for future studies. Am. J. Med. Genet. B Neuropsychiatr. Genet. 162B (8), 789–799.
Lunnon, K., et al., 2012. Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood. J. Alzheimers Dis. 30 (3), 685–710.
Lunnon, K., et al., 2014. Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. Nat. Neurosci. 17 (9), 1164–1170.
Lunnon, K., et al., 2016. Variation in 5-hydroxymethylcytosine across human cortex and cerebellum. Genome Biol. 17, 27.
Lutz, M.W., et al., 2010. Genetic variation at a single locus and age of onset for Alzheimer's disease. Alzheimers Dement. 6 (2), 125–131.
Lyko, F., et al., 2010. The honey bee epigenomes: differential methylation of brain DNA in queens and workers. PLoS Biol. 8 (11), e1000506.
Marambaud, P., et al., 2002. A presenilin-1/γ-secretase cleavage releases the E-cadherin intracellular domain and regulates disassembly of adherens junctions. EMBO J. 21 (8), 1948–1956.
Martinelli, P., Rugarli, E.I., 2010. Emerging roles of mitochondrial proteases in neurodegeneration. Biochim. Biophys. Acta Biomembr. 1797 (1), 1–10.
Masliah, E., et al., 2001. Altered expression of synaptic proteins occurs early during progression of Alzheimer's disease. Neurology 56 (1), 127–129.
Mattson, M.P., 2004. Pathways towards and away from Alzheimer's disease. Nature 430 (7000), 631–639.
McIlwain, D.R., et al., 2012. iRhom2 regulation of TACE controls TNF-mediated protection against listeria and responses to LPS. Science 335 (6065), 229–232.
Meda, S.A., et al., 2012. A large scale multivariate parallel ICA method reveals novel imaging-genetic relationships for Alzheimer's disease in the ADNI cohort. Neuro-Image 60 (3), 1608–1621.
Mi, H., et al., 2013. Large-scale gene function analysis with the PANTHER classification system. Nat. Protoc. 8 (8), 1551–1566.
Monson, N.L., et al., 2014. Elevated CNS inflammation in patients with preclinical Alzheimer's disease. J. Cereb. Blood Flow Metab. 34 (1), 30–33.
Naj, A.C., et al., 2010. Dementia revealed: novel chromosome 6 locus for late-onset Alzheimer disease provides genetic evidence for folate-pathway abnormalities. PLoS Genet. 6 (9), e1001130.
Naj, A.C., et al., 2011. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat. Genet. 43 (5), 436–441.
N'Diaye, E.N., et al., 2009. TREM-2 (triggering receptor expressed on myeloid cells 2) is a phagocytic receptor for bacteria. J. Cell Biol. 184 (2), 215–223.
Nelson, P.T., et al., 2014. ABCC9 gene polymorphism is associated with hippocampal sclerosis of aging pathology. Acta Neuropathol. 127 (6), 825–843.
Nestor, C.E., et al., 2012. Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. Genome Res. 22 (3), 467–477.
Neumann, H., Daly, M.J., 2012. Variant TREM2 as risk factor for Alzheimer's disease. N. Engl. J. Med.
Pan, X.D., et al., 2011. Microglial phagocytosis induced by fibrillar beta-amyloid is attenuated by oligomeric beta-amyloid: implications for Alzheimer's disease. Mol. Neurodegener. 6, 45.
Perez-Palma, E., et al., 2014. Overrepresentation of glutamate signaling in Alzheimer's disease: network-based pathway enrichment using meta-analysis of genome-wide association studies. PLoS One 9 (4), e95413.
Potkin, S.G., et al., 2009. Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. PLoS One 4 (8), e6501.
Pottier, C., et al., 2012. High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. Mol. Psychiatry 17 (9), 875–879.
Prince, M., Prina, M., Alzheimer's Disease International, 2013. Policy Brief for Heads of Government: the Global Impact of Dementia 2013–2050. Alzheimer's Disease International, London, pp. 1–8.
Proitsi, P., et al., 2014. Alzheimer's disease susceptibility variants in the MS4A6A gene are associated with altered levels of MS4A6A expression in blood. Neurobiol. Aging 35 (2), 279–290.
Puglielli, L., Tanzi, R.E., Kovacs, D.M., 2003. Alzheimer's disease: the cholesterol connection. Nat. Neurosci. 6 (4), 345–351.
Ramanan, V.K., et al., 2014. APOE and BCHE as modulators of cerebral amyloid deposition: a florbetapir PET genome-wide association study. Mol. Psychiatry 19 (3), 351–357.
Reiman, E.M., et al., 2007. GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. Neuron 54 (5), 713–720.
Ridge, P.G., et al., 2013. Alzheimer's disease: analyzing the missing heritability. PLoS One 8 (11), e79771.
Roses, M.D., Allen, D., 1996. Apolipoprotein E alleles as risk factors in Alzheimer's disease. Annu. Rev. Med. 47 (1), 387–400.
Roses, A.D., et al., 2009. Apoeε3 A3nd Tomm-40 haplotypes determine inheritance of Alzheimer's disease independently of Apoeε4 risk. Alzheimers Dement. 5 (4), e1.
Sakamuro, D., et al., 1996. BIN1 is a novel MYC-interacting protein with features of a tumour suppressor. Nat. Genet. 14 (1), 69–77.
Salminen, A., et al., 2009. Inflammation in Alzheimer's disease: amyloid-β oligomers trigger innate immunity defence via pattern recognition receptors. Prog. Neurobiol. 87 (3), 181–194.
Serrano-Pozo, A., et al., 2011. Neuropathological alterations in Alzheimer disease. Cold Spring Harb. Perspect. Biol. Med. 1 (1), a006189.

Seshadri, S., et al., 2010. Genome-wide analysis of genetic loci associated with Alzheimer disease. JAMA 303 (18), 1832–1840.
Sheng, J.G., Mrak, R.E., Griffin, W.S., 1997. Neuritic plaque evolution in Alzheimer's disease is accompanied by transition of activated microglia from primed to enlarged to phagocytic forms. Acta Neuropathol. 94 (1), 1–5.
Siggs, O.M., et al., 2012. iRhom2 is required for the secretion of mouse TNFalpha. Blood 119 (24), 5769–5771.
Slooter, A.J., et al., 1998. Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam study. Arch. Neurol. 55 (7), 964–968.
Song, C.X., et al., 2011. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. Nat. Biotechnol. 29 (1), 68–72.
Strittmatter, W.J., et al., 1993. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. Proc. Natl. Acad. Sci. U. S. A. 90 (5), 1977–1981.
Takahashi, K., Rochford, C.D., Neumann, H., 2005. Clearance of apoptotic neurons without inflammation by microglial triggering receptor expressed on myeloid cells-2. J. Exp. Med. 201 (4), 647–657.
Takei, N., et al., 2009. Genetic association study on in and around the APOE in late-onset Alzheimer disease in Japanese. Genomics 93 (5), 441–448.
Tuppo, E.E., Arias, H.R., 2005. The role of inflammation in Alzheimer's disease. Int. J. Biochem. Cell Biol. 37 (2), 289–305.
Webster, J.A., et al., 2008. Sorl1 as an Alzheimer's disease predisposition gene? Neurodegener. Dis. 5 (2), 60–64.
Wechsler-Reya, R., et al., 1997. Structural analysis of the human BIN1 gene. Evidence for tissue-specific transcriptional regulation and alternate RNA splicing. J. Biol. Chem. 272 (50), 31453–31458.
Wigge, P., McMahon, H.T., 1998. The amphiphysin family of proteins and their role in endocytosis at the synapse. Trends Neurosci. 21 (8), 339–344.
Wigge, P., et al., 1997. Amphiphysin heterodimers: potential role in clathrin-mediated endocytosis. Mol. Biol. Cell 8 (10), 2003–2015.
Wijsman, E.M., et al., 2011. Genome-wide association of familial late-onset Alzheimer's disease replicates BIN1 and CLU and nominates CUGBP2 in interaction with APOE. PLoS Genet. 7 (2), e1001308.
Wimo, A., 2010. World Alzheimer Report 2010; the Global Economic Impact of Dementia. In: A.s.D. International (Ed.), The Global Economic Impact of Dementia (London).
Yang, S., et al., 2008. Comparative proteomic analysis of brains of naturally aging mice. Neuroscience 154 (3), 1107–1120.
Yu, C.E., et al., 2007. Comprehensive analysis of APOE and selected proximate markers for late-onset Alzheimer's disease: patterns of linkage disequilibrium and disease/marker association. Genomics 89 (6), 655–665.
Yu, L., et al., 2014. Association of brain DNA methylation in SORL1, ABCA7, HLA-DRB5, SLC24A4, and BIN1 with pathological diagnosis of Alzheimer disease. JAMA Neurol.
Zhang, H., et al., 2005. Clusterin inhibits apoptosis by interacting with activated Bax. Nat. Cell Biol. 7 (9), 909–915.

APPENDIX 2 – POST-MORTEM BRAIN TISSUE AND ITS APPLICATION TO STUDY EPIGENETIC REGULATION IN ALZHEIMER'S DISEASE

POST-MORTEM BRAIN TISSUE AND ITS APPLICATION TO STUDY EPIGENETIC REGULATION IN ALZHEIMER'S DISEASE

Adam R Smith[1] and Katie Lunnon[1,*]

[1.] University of Exeter Medical School, RILD Level 4, Barrack Road, Exeter, Devon, UK.

[*] Corresponding author at: University of Exeter Medical School, RILD, Barrack Road, University of Exeter, Devon, UK. UK. Tel: + 44 1392 408 298 Email address: k.lunnon@exeter.ac.uk

ABSTRACT

The study of epigenetic mechanisms in many complex disease phenotypes has recently gained considerable research interest. In recent years a number of epigenome-wide association studies (EWAS) of DNA methylation have been published in a range of brain disorders, including Alzheimer's disease (AD). These have utilized post-mortem human brain tissue to investigate, in detail, epigenetic changes in a disease relevant tissue from the organism of interest. However epigenetic studies of brain disorders present specific experimental design challenges. Here, we describe selected experimental designs that have used post-mortem brain tissue to investigate and evaluate DNA methylation in AD and other brain disorders. We also evaluate a new approach to analyze DNA hydroxymethylation, a recently identified epigenetic modification. Ultimately the use of this method in AD will provide the field with a more in depth understanding of genomic regulation in disease. Our aim is to provide guidelines and advice for researchers contemplating undertaking EWAS in AD and other brain disorders, from the isolation of DNA to performing EWAS in post-mortem brain tissue.

1. INTRODUCTION

1.1 Moving beyond genetic variation to understand Alzheimer's disease: The advent of epigenome wide association studies

Genomic studies for various brain diseases, including Alzheimer's disease (AD) are now commonplace in scientific literature[1-10]. Genome wide association studies (GWAS), in particular, having yielded a vast amount of the current knowledge for genetic risk factors for AD. However, for many brain disorders GWAS have reached a saturation point with respect to their ability to identify novel genetic variants[11,12]. Furthermore although whole genome and exome sequencing studies are beginning to uncover additional polymorphisms associated with disease, these newly discovered variants are usually relatively rare within the population[13]. Therefore, recent studies have extended to investigating genomic regulation via epigenetic mechanisms to explain the sporadic, yet partially heritable nature of several diseases, including AD. The term "epigenetics" refers to the reversible regulation of gene expression mediated principally through changes in DNA methylation and chromatin structure, occurring independently of the DNA sequence.[14] To date, epigenome wide association studies (EWAS) for brain disorders have predominantly focused on studying DNA methylation patterns. There are however several additional study design challenges compared to traditional GWAS, for example careful consideration of methodological approaches, tissue-specificity and causal inference[15].

Although DNA methylation (5-methylcytosine (5-mC)) is the best understood and most studied epigenetic modification, additional DNA modifications have been recently described, for example 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC). These modifications were originally thought to represent intermediates in the demethylation of 5-mC to un-modified cytosine[16], however recent data suggests that 5-hmC

may also be a functional epigenetic mark, showing tissue-specific patterns in the brain[17]. DNA methylation and hydroxymethylation of cytosine results in the re-modeling of the 3-dimensional structure of chromatin, either preventing, or enabling the binding of transcription factors or splicing machinery that results in changes in the transcriptional properties of genes as well as the splice variants that translate into protein products[18]. This change in the architecture of the gene/protein expression landscape has been hypothesized to be the cause of many diseases that currently are without a definitive genetic cause, such as AD. However, it is worth noting that epigenetic changes occur at a cellular level, with DNA methylation patterns occurring in a more disparate manner between tissues of the same individual, than between individuals within the same tissue[19]. In AD, the progression of disease is well characterized, with the spread of hyperphosphorylated tau being predictable in its spread throughout the brain, starting in the brain stem and the entorhinal cortex, moving through the prefrontal cortex and the limbic regions and in late Braak stage, being present throughout the neocortex, with the cerebellum remaining relatively free of neuropathology[20]. As such, examining post-mortem brain tissue is essential for epigenetic studies in AD. However, it is important to choose an appropriate tissue that accurately represents the pathology of AD. Looking in areas of the brain that are affected early in AD mean that those who have a clinical diagnosis of AD will already have large amounts of neuronal loss and so being able to assess epigenetic changes in affected neurons is more challenging. Using brain regions affected later in the disease means fewer people will have pathology in these regions and so may be an indication of what is changing during the spread of pathology. A cross-tissue study design is therefore optimal as it allows the investigation of multiple different brain regions representing the spectrum of pathology. As the epigenetic field moves forward it will be important to integrate multi-omic approaches (genomics, epigenomics, and transcriptomics), to fully elucidate molecular changes occurring during disease pathology.

1.2 Methods to study DNA methylation and DNA hydroxymethylation profiles in Alzheimer's disease

At present epigenomic studies in AD represent a relatively new scientific endeavor, with considerably more review articles than empirical research. Of the studies that have looked at DNA methylation profiles in AD brain samples, most methods have used a targeted approach to determine the DNA methylation profile in specific genes of interest[21] or used immunohistochemical methodology to investigate the total global levels of 5-mC and or 5-hmC in disease[22-24]. However, in recent years the first EWAS in AD have been published, with six studies performed to date[25-30], although only two had sufficient sample numbers to identify experiment-wide significant results[25,26]. In this chapter we will focus on describing methods to perform EWAS to characterize 5-mC and 5-hmC profiles in post mortem brain tissue. In addition to presenting a method to isolate sufficient yields of high molecular weight DNA, we will also describe the EWAS approach taken by Lunnon *et al* and De Jager *et al* in 2014[25,26], which utilized the Illumina Infinium 450K BeadChip array in conjunction with bisulfite-treated DNA to assess DNA modifications (sum of 5-mC and 5-hmC) at ~485,000 CpG sites in the genome. Although this method has since been superseded by the Illumina Infinium EPIC BeadChip array, which quantifies DNA modifications at ~850,000 CpG sites in the genome, the principles are essentially the same. We also describe a new methodology that has recently become available[17], which involves bisulfite and oxidative bisulfite treatment of DNA in parallel, in conjunction with the Illumina Infinium BeadChip array to assess 5-mC and 5-hmC levels separately in the genome. Although this method has been used to study 5-mC and 5-hmC levels in post-mortem brain samples, it has not, to date, been applied to the study of AD. We will describe the practical elements of the appropriate study design, discussing optimal methods for DNA extraction, DNA bisulfite and oxidative bisulfite treatment and finally DNA methylation profiling on the Illumina Infinium BeadChip arrays. When performing the described experiments

it is important, at all times, to be hypervigilant to contamination between samples and from external sources.

2. MATERIALS AND METHODS

2.1 DNA extraction from post mortem tissue

There are many methods and protocols available to isolate DNA from human tissue samples. DNA extraction from brain tissue samples using a standard phenol-chloroform method has been widely utilized for EWAS using post-mortem brain tissue as it provides sufficient yields of high quality DNA for downstream purposes. The protocol below is an adaptation of the phenol-chloroform method developed by Sambrook and Russell (2006)[31]. Reagents required for the Phenol-Chloroform DNA extraction can be found in **Tables 1.1, 1.2** and **1.3**.

2.1.1 Weigh brain tissue

Perform steps 1-6 in a sterile extraction hood (see **Note 1.1**)

1. Weigh empty frozen 1.5ml microcentrifuge tube and note the weight
2. Transfer frozen tissue sample to the 1.5ml microcentrifuge tube with sterile tip or tweezers whilst on dry ice
3. Weigh the microcentrifuge tube and calculate the brain weight (total weight-empty weight). The optimum protocol utilizes 100mg of tissue (see **Note 1.2**)

2.1.2 Lyse and digest cells

4. Prepare sufficient fresh lysis buffer (1ml per 100mg tissue sample) (**Table 1.1**)

5. Add 4 times weight/volume of lysis buffer, for example if utilizing 100mg brain tissue add 400µl lysis buffer
6. Disrupt the tissue with a disposable pestle and rotary pestle motor. Subsequently discard the pestle
7. Add 6 times weight/volume of lysis buffer (for example if utilizing 100mg brain tissue add 600µl lysis buffer) (see **Note 1.3**)
8. Add DNase-free RNase-A (**Table 1.2**) to a final concentration of 20µg/ml and incubate for at least 1hr at 37°C (see **Note 1.4**)
9. Add Proteinase K (**Table 1.2**) to each sample at a final concentration of 400µg/ml
10. Incubate in a water bath at 50°C overnight (see **Note 1.5)**
11. The following morning check the level of sample lysis. If the sample is not sufficiently lysed use a p1000 pipette tip or a syringe to re-suspend the solution and break up any remaining tissue, being careful not to lose any of the solution. In most cases lysis has been achieved and this step is not necessary (see **Note 1.6**)
12. Add Proteinase K (**Table 1.2**) to each sample at a final concentration of 200µg/ml
13. Leave for at least one more hour in the water bath at 50°C
14. Heat the water bath to 65°C and leave for 20-30 minutes to deactivate the Proteinase K
15. Cool the sample to room temperature

2.1.3 Purification of nucleic acids (Perform these steps in a chemical hood: see **Note 1.7**)

16. Transfer sample to a 2ml phase lock tube and then add an equal volume of Phenol / Chloroform / Isoamyl Alcohol (PCI) to the tube (**Table 1.2**) (for example, add 1000µl if utilizing 100mg brain tissue) (see **Note 1.8)**
17. Invert sample 20 times to mix
18. Centrifuge 16200 x g for 15 minutes
19. Carefully remove the aqueous layer into a new 2ml microcentrifuge tube (see **Note 1.9**)

20. Add 1ml of 100% chloroform (**Table 1.2**)
21. Invert sample 20 times to mix
22. Centrifuge 16200 x g for 15 minutes
23. Remove the aqueous layer into a new tube 2mL microcentrifuge tube (see **Note 1.10**)
24. Repeat chloroform extraction steps 20 to 23

2.1.4 Precipitation of DNA

25. Add 1 to 1.5ml of ice cold 100% ethanol (**Table 1.2**) to the sample (see **Note 1.11**)
26. Mix the sample slowly (see **Note 1.12**)
27. If DNA can not be visualized, add between 10-20µl 5M NaCl (**Table 1.3**)
28. Leave at -20°C for one hour to further precipitate (see **Note 1.13**)
29. Centrifuge at 16200 x g for 15 minutes
30. Remove supernatant by pipetting, being careful not to dislodge the DNA pellet
31. Add 1ml 70% ethanol (**Table 1.2**)
32. Centrifuge at 16200 x g for 5 minutes, or if pellet detaches from the base of the tube for at least 10 minutes
33. Remove the supernatant by pipetting, being careful not to dislodge the DNA pellet. Leave the pellet to air-dry for approximately 30 minutes (see **Note 1.14**)
34. Re-suspend in 200µl - 300µl H_2O or Te (**Table 1.3**) (see **Note 1.15**)
35. Dissolve the pellet overnight at 37°C

2.1.5 DNA quality check

36. Perform a visual check to make sure the DNA is well suspended (see **Note 1.16**)
37. Run DNA on a 0.8% agarose gel to check for high molecular weight DNA, which is not degraded.

38. Load 2µl of the DNA sample on to a Spectrometer (NanoDrop) to estimate the concentration and further check the quality using 260/280 and 230/280 ratios (above 1.8 and ~2 respectively for pure DNA)

2.2. Bisulfite treatment of DNA

Bisulfite treatment, originally described by Frommer *et al*[32], is a method that provides a base-pair level sensitivity measure for DNA modifications (including 5-mC). In this process un-modified cytosine is converted to uracil, but certain modifications to cytosine including 5-mC and 5-hmC remains unreactive. The sequence can then be amplified via subsequent PCR, resulting in all uracil and thymine residues being amplified as thymine and only modified cytosine residues being amplified as cytosine. Bisulfite treatment using the Zymo EZ-96 DNA Methylation-Gold™ Kit ideally utilizes 500ng high quality DNA in 20µl to provide sufficient bisulfite treated material for cytosine modification array profiling and subsequent validation experiments. This method has been used to treat the DNA samples in all EWAS studies of AD published to date[25-30]. The steps for the protocol are described below, and the manufacturer's protocol is available from http://www.zymoresearch.com/downloads/dl/file/id/59/d5007i.pdf. Before you start the protocol make a dilution plate of DNA at 25ng/µl in 20µl in the conversion plate provided with the kit.

1. Add 130µl of the CT conversion reagent (**Table 2.1**) to each DNA sample in the conversion plate and mix the samples by pipetting
2. Seal the plate with the provided film. Transfer the conversion plate to a thermal cycler and perform the following steps:
    1. 98°C for 10min
    2. 64°C for 2.5 hours
    3. 4°C storage for up to 20 hours

3. Add 400µl of M-Binding buffer (provided) to the wells of a Silicon-A™ binding plate (provided) mounted on a collection plate (provided)
4. Transfer the samples from the conversion plate (Step 2) to the wells of the Silicon-A™ binding plate (Step 3). Mix by pipetting up and down
5. Centrifuge at 3500 x g for 5 minutes. Discard the flow-through (see **Note 2.2**)
6. Add 400 µl of M-Wash buffer (**Table 2.2**) to each well of the plate
7. Centrifuge at 3500 x g for 5 minutes
8. Add 200 µl of M-Desulphonation buffer (provided) to each well and allow the plate to stand at room temperature (20-30°C) for 20 minutes
9. Centrifuge at 3500 x g for 5 minutes. Discard the flow-through
10. Add 400 µl of M-Wash buffer to each well of the plate
11. Centrifuge at 3500 x g for 5 minutes. Discard the flow-through
12. Add 400 µl of M-Wash buffer (**Table 2.2**) and centrifuge for 10 minutes
13. Place the Silicon-A™ binding plate onto an elution plate (provided)
14. Add 15 µl of M-Elution buffer (provided) directly to each well of the Silicon-A™ binding plate. (see **Note 2.3**)
15. Incubate for 5 minutes at room temperature
16. Centrifuge at 4000 x g for 3 minutes to elute the DNA
17. Repeat steps 15 to 17.

Eluted DNA is ready for immediate analysis or can be stored at or below -20°C for later use. For long term storage, store at or below -70°C. To assess bisulfite conversion quality it is recommended to perform a bisulfite-specific PCR amplification (details available on request) followed by gel electrophoresis. If a PCR product is visible this is a good indication of bisulfite conversion quality.

2.3 Oxidative bisulfite treatment of DNA

Oxidative bisulfite treatment is the first method enabling the quantitative mapping of 5-mC in the absence of confounding by 5-hmC in genomic DNA at single-nucleotide resolution. Using a selective chemical oxidation of 5-hmC to 5-fC, followed by bisulfite conversion, results in the conversion of 5-hmC to uracil. As such only cytosines with DNA methylation are protected from the conversion and after subsequent PCR amplification, one can assess the level of just 5-mC. Simultaneous bisulfite treatment of a matched sample of DNA, provides a value for the total DNA modifications (5-mC + 5-hmC) at each site, as both modifications are protected from the conversion and are not converted to uracil[33]. Subsequently after PCR amplification, the subtraction of the oxidative bisulfite signal from the standard bisulfite signal gives a 5-hmC proxy value for each nucleotide.

Oxidative bisulfite treatment using the CEGX® TrueMethyl® Array kit requires 1µg high molecular weight genomic DNA per sample. This is sufficient to allow 500ng to be processed through each of the bisulfite and oxidative bisulfite conversion processes. Input masses of high molecular weight gDNA in the range of 50ng – 1µg are also compatible (see **Note 3.1**). The protocol described results in a final volume of 12µl of bisulfite treated DNA and 12µl of oxidative bisulfite treated DNA for downstream analyses. The manufacturer's protocol is available from https://www.cambridge-epigenetix.com/resources/user-guides

2.3.1 Starting material

1. Add 1µg DNA sample to a 1.5ml microcentrifuge tube and adjust the final volume to 50µl using ultra-pure H_2O from the TrueMethyl® kit (see **Note 3.2**)

2.3.2 DNA oxidation

2. Set heat block to 37°C and prepare Magnetic binding solution 1 (**Table 3.1**) and 80% acetonitrile (**Table 3.2**) (see **Note 3.3**)

3. Vortex Magnetic bead binding solution 1 thoroughly to ensure the solution is homogenous and add 100µl of Magnetic bead binding solution 1 to each 1.5ml microcentrifuge tube containing 50µl DNA and vortex to mix (see **Note 3.4**)
4. Incubate the sample for 20 minutes at room temperature
5. At the end of the incubation, briefly centrifuge the 1.5ml microcentrifuge tubes to collect the sample at the bottom of the tube
6. Place the tubes into a magnetic separation rack and pellet beads for 5 minutes at room temperature
7. Carefully remove and discard the supernatant, avoiding the bead pellet. If beads are accidentally pipetted, dispense them back into the tube and repeat step 6
8. Leaving the 1.5ml tubes in the magnetic separation rack, complete the following wash steps: -
    a. Carefully add 1ml 80% Acetonitrile (**Table 3.2**) wash to the tubes without disturbing the bead pellet
    b. Remove and discard the 1ml 80% acetonitrile wash, carefully avoiding aspiration of the bead pellet
    c. Repeat Steps 8a and 8b another two times so that ultimately three 1ml 80% acetonitrile washes will have been performed in total. Remove, by pipetting, as much of the final wash as possible
    d. Air dry the bead pellets for 5 minutes at room temperature, leaving the lids of the tubes open
9. Complete the following denaturation steps: -
    a. Leaving the 1.5ml tubes in the magnetic separation rack, add 50µl of Denaturing solution (provided) directly onto the bead pellet. Close the lid of the tube before removing from the magnetic separation rack
    b. Vortex thoroughly to fully suspend the bead pellet in the Denaturing solution

c. Incubate for 5 minutes at room temperature
d. Incubate for 30 minutes at 37°C (perform this step in the pre-heated benchtop heating block (step 2))

10. Briefly centrifuge the 1.5ml microcentrifuge tubes to collect the sample at the bottom of the tube
11. Place the tubes into a magnetic separation rack and pellet beads for 5 minutes at room temperature
12. Avoiding the bead pellet, carefully transfer 50µl eluate into a new 1.5ml tube. Centrifuge briefly. Place sample on ice and proceed immediately to step 13
13. Perform the DNA oxidation reaction in 0.2ml PCR tubes as described in **Table 3.3**.
14. Incubate the oxidation reaction mix at 40°C for 30 minutes in a PCR thermocycler with a heated lid (see **Note 3.5**)
15. Centrifuge the sample at 14000 x g for 10 minutes at room temperature to pellet any black precipitate. If the supernatant remains orange, there is an excess of oxidant remaining and oxidation has completed (see **Note 3.6**)
16. Transfer the orange supernatant into a new 0.2ml PCR tubes and use this in subsequent steps of the protocol. Take care not to carry any black precipitate over as this could inhibit downstream steps. Proceed immediately to the bisulfite conversion steps described below

2.3.3 Bisulfite conversion

17. Prepare the Bisulfite reagent solution by adding 4.2ml Bisulfite diluent to one Bisulfite reagent aliquot (sufficient for 24 reactions). Seal the lid tightly and incubate the solution at 60°C for 15 minutes (using either a water bath or heated orbital incubator).

Subsequently vortex the sample until the Bisulfite reagent solution is completely dissolved (see **Note 3.7**)

18. Equilibrate the oxidized samples from step 16 to room temperature. Do not place the oxidized samples on ice to cool as this may cause the solution to precipitate
19. Perform the bisulfite conversion reactions in the same 0.2ml PCR tubes (**Table 3.4**).
20. Place the 0.2ml PCR tube containing the Bisulfite conversion reaction mix into a thermocycler. Perform the bisulfite DNA conversion using the thermal profile shown in **Table 3.5**. The complete cycle should take approximately 8 hours to complete. This step may be performed overnight if necessary (see **Note 3.8**)

2.3.4 Clean up of bisulfite converted DNA

21. Prepare Magnetic binding solution 2 (**Table 3.6**)(see **Note 3.9**)
22. Equilibrate Magnetic bead Binding solution 2 and Desulfonation buffer (**Table 3.7**) to room temperature for a minimum of 30 minutes before use. Also prepare a fresh stock of 70% ethanol for the experiment (**Table 3.8**)
23. Once the bisulfite conversion is complete, centrifuge the 0.2ml PCR tubes containing the reactions at 14000 x g for 10 minutes at room temperature. This will pellet any precipitated salts accumulated during the bisulfite incubation
24. Without disturbing the pellet, transfer 195µl of the supernatant to a new 1.5ml microcentrifuge tube (see **Note 3.10**)
25. Vortex Magnetic bead binding solution 2 (**Table 3.6**) thoroughly to ensure the solution is homogenous before aliquoting
26. Carefully add 1ml of Magnetic bead binding solution 2 (**Table 3.6**) to each of the 1.5ml microcentrifuge tubes containing 195µl bisulfite converted sample from step 24. Close the lid of the tube. Invert tube and flick the base of the tube to dislodge the sample,

vortex immediately to mix. Place the tube at a 45 degree angle to ensure thorough mixing

27. Incubate for 20 minutes at room temperature (see **Note 3.11**)
28. At the end of the incubation, briefly centrifuge the 1.5ml microcentrifuge tubes to collect the sample at the bottom of the tube
    a. Place the tubes into a magnetic separation rack and pellet beads for 30 minutes at room temperature (see **Note 3.12**)
    b. Carefully remove and discard the supernatant, avoiding the bead pellet. If beads are accidentally pipetted, dispense them back into the tube and repeat step 28a
29. Leaving the 1.5ml tubes in the magnetic separation rack, complete the following wash steps:
    a. Add 1.4ml 70% ethanol wash (**Table 3.8**) to the tubes without disturbing the bead pellet
    b. Remove and discard the 1.4ml 70% ethanol wash, carefully avoiding aspiration of the bead pellet. Remove by pipetting as much of the wash as possible
30. Complete the following desulfonation steps:
    a. Leaving the 1.5ml tubes in the magnetic separation rack, add 1ml Desulfonation buffer (**Table 3.7**) directly onto the bead pellet. Close lid of tube before removing from the magnetic separation rack. Vortex to fully suspend the bead pellet in Desulfonation buffer and centrifuge briefly
    b. Place the tubes back into the magnetic separation rack and incubate for 5 minutes at room temperature in the Desulfonation buffer. The beads will pellet during this incubation step
    c. Remove and discard the 1ml Desulfonation buffer, carefully avoiding aspiration of the bead pellet. Remove by pipetting as much of the Desulfonation buffer as possible

31. Leaving the 1.5ml tubes in the magnetic separation rack, complete the following wash steps:
    a. Add 1ml 70% ethanol wash to the tubes without disturbing the bead pellet.
    b. Remove by pipetting and discard the 1ml 70% ethanol wash, carefully avoiding aspiration of the bead pellet.
    c. Repeat Steps 31a and 31b so that ultimately 1ml 70% ethanol washes will have been performed two times in total. Remove by pipetting as much of the final wash as possible
    d. Air dry the bead pellets for 15 minutes at room temperature, leaving the lids of the tubes open
32. Complete the following elution steps:
    a. Leaving the 1.5ml tubes in the magnetic separation rack, add 12µl of Elution buffer (provided) directly onto the bead pellet. Close lid of tube before removing from the magnetic separation rack.
    b. Vortex thoroughly to fully suspend the bead pellet in the Elution buffer. Incubate for 20 minutes at room temperature to elute the TrueMethyl® converted DNA from the beads.
33. At the end of the incubation (step 32b), briefly centrifuge the 1.5ml microcentrifuge tubes to collect the sample at the bottom of the tube. Place the tubes into a magnetic separation rack and pellet beads for 5 minutes at room temperature (see **Note 3.13**)
34. Avoiding the bead pellet, carefully transfer 10 - 12µl eluate into a new 1.5ml tube. This is the recovered TrueMethyl® converted DNA sample

The DNA is ready for immediate analysis or can be stored at or below -20°C for later use. For long term storage, store at or below -70°C. To assess oxidative bisulfite conversion quality it is recommended to integrate a Digestion control (provided) with the initial 1µg gDNA sample in

step 1 (see **Note 3.2**) and interrogate the control using PCR, followed by gel electrophoresis (see also the TrueMethyl® Array User guide).

3.3 Illumina Infinium Methylation BeadChip Array

The Infinium HD Methylation Assay combines bisulfite converted DNA and whole-genome amplification with direct, array-based capture and scoring of the CpG loci[34]. Signal intensity is measured with the Illumina iScan or HiScan system to generate beta values, a measure of the degree of 5-mC/5-hmC at each locus. The beta values can then be interrogated and compared across samples for powerful large-scale studies. One or two probes are used to interrogate a CpG locus, depending on the probe design for a particular CpG site. (The Infinium I design has two probes per site and Infinium II has one probe per site.) Allele-specific single base extension of the probes incorporates a biotin nucleotide or a dinitrophenyl labeled nucleotide. Signal amplification of the incorporated label further improves the overall signal-to-noise ratio of the assay.

The first stage in the process includes a denaturing and neutralization stage before the bisulfite treated DNA is whole-genome amplified (see **Note 4.1**). The denatured DNA is isothermally amplified in an overnight step. The whole-genome amplification uniformly increases the amount of the DNA sample by several thousand folds without significant amplification bias. A controlled enzymatic process then fragments the amplified product. The process uses endpoint fragmentation to prevent over fragmentation. After an isopropanol precipitation, centrifugation at 4°C collects the fragmented DNA. The precipitated DNA is then suspended in Hybridization buffer. DNA samples are then dispensed onto BeadChip arrays and incubated in an Illumina Hybridization Oven to hybridize the samples onto the BeadChip array surface. The new EPIC array design enables eight samples to be applied to each BeadChip array, which keeps them

separate with an IntelliHyb seal. DNA samples anneal to locus-specific sites on the BeadChip array. Un-hybridized and nonspecifically hybridized DNA are washed away and the BeadChip array then undergoes extension and staining in capillary flow-through chambers. Single-base extension of the oligonucleotides on the BeadChip array, using the captured DNA as a template, incorporates detectable labels on the BeadChip array and determines the 5-mC/5-hmC level of the specific CpG sites. Finally the BeadChip array is imaged on the Illumina HiScan or iScan System, using a laser to excite the fluorophore of the single-base extension product on the beads. The scanner records high resolution images of the light emitted from the fluorophores. These in turn are used by the genome studio® software to produce a beta value at each locus. A full description of this methodology can be found on the Illumina website http://www.illumina.com/products/by-type/microarray-kits/infinium-methylation-epic.html.

The Infinium HD Methylation Assay platform has opened up multiple possibilities for epigenomic research that were previously not feasible; analyzing over 850,000 CpG positions (EPIC array) in the human genome has enabled researchers to gain a snap-shot of the global modification profile across multiple samples. This approach is considerably more affordable than current sequencing alternatives, however the cost per sample is still relatively high at approximately £250 per sample. Initially designed as a cancer profiling tool, a large percentage of sites on the array are tailored to cancer related targets, a factor that needs to be considered when studying other phenotypes. In addition the HD Methylation BeadChip array contains multiple probes where cross-hybridizing with multiple DNA fragments is common, coupled with a number of probes that contain common small nucleotide polymorphisms (SNPs); both of which should be excluded from downstream analysis. As a result of this, and the large amount of data generated by this technique, specialist statistical pipelines and digital storage facilities for QC and data analysis are required.

4. NOTES

4.1 Notes for DNA extraction from post mortem brain tissue

Note 1.1 It is important when undertaking DNA extraction to be hyper-vigilant to contamination between samples and from external sources. Materials should be clean-room quality and autoclaved where possible or purchased sterile/DNase-free. A sterile technique should be used, and gloves should be changed regularly.

Brain tissue may carry pathogenic material (for example from donors with infectious diseases). Therefore performing the first stages in a chemical hood provides the operator with an extra level of protection. The quality and ease of this extraction procedure is highly dependent on the starting material, for example a 100mg sample of cortical tissue will give approximately half the amount of DNA to an identically sized sample of cerebellum tissue. However, donor cerebellum samples usually yield higher levels of contaminates in the final DNA sample, compared to other brain regions. The level of impurities seen in DNA from the cerebellum does not appear to affect downstream analysis of the DNA methylome; however if other analytical methods are going to be performed, it can be useful to perform a second PCI cleaning stage.

Note 1.2 When dissecting small amounts of brain tissue it is important to keep all equipment that is in contact with the tissue sterile and cold, for example keeping the scalpel, microcentrifuge tube and petri dish on dry ice. This prevents the tissue thawing as it is being weighed. In addition, condensation does appear on the surface of chilled microcentrifuge tubes, which may have an effect on the weight of the tube, to avoid this quickly rub off condensation with some clean tissue before weighing the tube for the first time. If you plan to extract DNA from large numbers of samples, it is advisable to cut all tissues to the recommended weight first and place back at -80°C. You can then return to these prepared tissues when you are ready to extract all of the DNA samples.

Note 1.3 This section results in the fast vortexing /mixing of the tissue and lysis buffer, if the volume of lysis buffer is too large, the tissue/buffer is more likely to spill out at this stage. Larger tubes are not advised as the rotary pestle recommended is designed to fit the shape of the 1.5ml microcentrifuge tube. The pestle can be used to grind the frozen tissue against the side of the tube before the rotary motor is used, when removing the pestle from the lysis buffer/tissue mix, be careful to remove as much tissue as possible from the pestle by scraping on the side of the 1.5ml tube. Used pestles should be disposed of in line with individual institution's recommendations for tissue-contaminated equipment.

Note 1.4 One hour is the minimum recommended time for RNase treatment; however it is better to incubate for approximately two hours. Adding the RNase at this stage eliminates the need to remove RNA from the semi-purified DNA at a later stage in preparation, pancreatic RNase is not highly active in the presence of 0.5% SDS, but when added at high concentrations, it is sufficiently active to degrade most of the cellular RNA. If RNA and DNA are needed for downstream experiments, we recommend using the All-prep DNA/RNA mini kit available from Qiagen.

Note 1.5 Proteinase K digests protein in the tissue to expose DNA; therefore it is not appropriate for the study of histone modifications. The optimum temperature for enzymatic activity is 50°C; the optimum time needed for tissue lysis is dependent on brain tissue.

Note 1.6 A good sign of lysis is a pale/transparent solution, if tissue is still too viscous to pipette use some sterile scissors to cut the tip off a p1000 pipette tip, this will make pipetting easier, pipette up and down slowly several times to shear the material further. This method can also be performed before the 50°C incubation with Proteinase K if the pestle was insufficient at breaking up all the brain tissue.

Note 1.7 Chloroform and PCI are harmful and there is a danger of serious damage to health by prolonged exposure through inhalation and swallowing. Wear protective clothing, including safety glasses at all times and performing these steps in the chemical hood is strongly advised.

Note 1.8 Phase lock tubes can be bought (MaXtract High Density, Qiagen). An alternative is to make in-house equivalents as follows:

Carefully add 1ml vacuum grease (Dow Corning® high-vacuum silicone grease (Sigma)) to the cap of a 2ml microcentrifuge tube, using a 50ml syringe (avoiding contamination of the syringe tip). Shut the lid and centrifuge at 1200 x g for 10 minutes before use.

Note 1.9 The aqueous layer contains genetic material, whist the separated (waste) layer contains degraded protein and other contaminates.

Note 1.10 A cloudy white precipitate may be visible on the surface of the aqueous layer. This layer of protein and/or detritus has not been efficiently removed by the phenol-chloroform extraction. Remove as much of the aqueous layer as possible while avoiding this white precipitate. The inclusion of small amounts of this precipitate in your final sample will not usually have any detrimental downstream effects on DNA quality or DNA methylation analysis.

Note 1.11 Cold 100% ethanol aids in the precipitation process of DNA. Store the ethanol at -20°C for approximately 30 minutes prior to use.

Note 1.12 At this stage a “cotton wool” like substance should become visible in the ethanol mix, which is the precipitated DNA. Using the visible size of the DNA at this point is a good indication of the amount of elution buffer (H_2O/Te) needed at the final stage (for example xmm diameter of pelleted DNA = x100µl elution buffer, to a maximum volume of 300µL).

Note 1.13 Low temperatures aid in the precipitation process, samples can be left overnight at -20°C. If processing large numbers of samples it is recommended to pause at this stage. If you

decide to incubate the extractions at -20°C overnight, it is important to perform this on all samples in the cohort to maintain consistency and minimize batch effects.

Note 1.14 Do not leave the samples air-drying too long, otherwise it will be difficult to re-suspend the DNA pellet. Once the smell of ethanol has disappeared and there are no visible drops of liquid on the inside of the tube it is appropriate to move on to the next stage.

Note 1.15 Some applications, for example oxidative bisulfite treatment, are very sensitive to the presence of alcohols (ethanol, isopropanol, phenol) as well as solutions containing specific compounds (Tris, EB buffer, TE buffer, glycerol, surfactants). If the use of the standard bisulfite treatment kit from Zymo is planned then Te is recommended as it provides better stability for long term DNA storage. For all other applications ultra-pure H_2O is recommended.

Note 1.16 A clear colorless solution shows the DNA is completely dissolved; the solution could still be very viscous due to the concentration of DNA. It is recommended to dilute each sample to 100ng/µl to allow accurate pipetting at later stages.

4.2 Notes for bisulfite treatment of DNA

Note 2.1 The CT conversion reagent is best dissolved by continuous shaking over 15 minutes. Securing the tube to a vortexer enables the preparation of other reagents, whilst the CT reagent is dissolving.

Note 2.2 Throughout all centrifugation steps the plates should not have a cover film applied as this will prevent efficient movement through the filter. The plates do become very full so care is needed to avoid spillage of samples into adjacent wells. Remove flow-through at each stage to prevent back-contamination of the filters from the wells of the collection plate.

Note 2.3 The standard manufacturer's protocol uses one volume of 30µl to elute DNA, here we have documented a change which we have found to yield greater amounts of converted DNA, without changing the final elution volume.

4.3 Notes for oxidative bisulfite treatment of DNA

Note 3.1 Processing samples through the TrueMethyl® Array workflow in batches of 12 to produce 12 oxidative bisulfite sample reactions and 12 bisulfite sample reactions is recommended. If higher throughput is required a 96 well plate format is available from CEGX®

Note 3.2 For integration of Digestion control at this stage see TrueMethyl® Array User guide step 2.1, followed by step 5 for interrogation of Digestion Control. Inclusion of these steps is highly recommended for the first run though of this protocol.

Note 3.3 Aliquoting the prepared Magnetic bead binding solution 1 into 8 x 1.5ml microcentrifuge tubes each containing 1.4ml per tube is recommended. One aliquot may then be used per batch of 12 samples processed.

Note 3.4 No more than 50µl of DNA should be added to the bead purification at this step, if the DNA concentration is too low a concentration step will need to be performed.

Note 3.5 The preparation of the bisulfite reagent solution (step 17) may be completed during this incubation time.

Note 3.6 If the solution turns any color other than orange after the 30 minute incubation, it is a sign that the oxidation reaction is compromised. This may lead to incomplete conversion of 5-hmC to uracil and variability in your results. To avoid this, ensure only ultra-pure H_2O provided with the TrueMethyl® kit is used in steps containing the oxidant and consider re-purification of

your sample taking all practicable steps to remove contaminants from the starting DNA sample solution.

Note 3.7 If processing more than one batch of 12 samples, ensure to use the same Bisulfite reagent solution aliquot for each oxidative bisulfite and bisulfite treated sample pair.

Note 3.8 If using a thermal cycler that does not allow you to enter the reaction volume (200µl) set the instrument to the largest volume setting available. Since the bisulfite reaction is not overlaid with mineral oil, only thermal cyclers with heated lids are suitable for this procedure. Converted DNA can be left in the thermal cycler overnight at 20°C hold without any loss of quality.

Note 3.9 Aliquoting the prepared Magnetic bead binding solution 2 into 8 x 50ml polypropylene centrifuge tubes each containing 27ml per tube is recommended. One aliquot may then be used per batch of 12 samples processed.

Note 3.10 Failure to perform this centrifugation step will cause compounds to be carried forward into the PCR reaction that will inhibit amplification. Use 0.2ml tube adapters in the centrifuge to perform this step.

Note 3.11 Magnetic bead solution 2 is a viscous solution and requires slow and careful pipetting to ensure 1ml is dispensed fully into the sample.

Note 3.12 At the end of the 30 minute incubation the supernatant should become clear. If the supernatant isn't clear after 30 minutes, pellet the beads for a longer period of time.

Note 3.13 Due to the small elution volume, the tubes should be placed into the rack in a slightly elevated position as shown in **Figure 1**. This helps to keep the bead pellet from separating in the tube, therefore preventing accidental bead uptake and ensuring efficient recovery of eluate. Placing a plastic tube rack upside down beneath the tubes is recommended to keep the 1.5ml microcentrifuge tubes in an elevated position.

3.4 Notes for Illumina Infinium Methylation BeadChip Array

Note 4.1 The Infinium methylation array has a comprehensive set of notes attached to its protocol for further information. One alteration to this standard protocol that is recommended is to increase the input volume of bisulfite or oxidative bisulfite treated DNA from 4µl to 7µl and simultaneously reduce the volume of NaOH whilst maintaining final concentration (1µl, 0.4M NaOH). This increase in DNA volume increases the success rate of the array process without changing the overall volume of the whole-genome amplification.

5. Conclusion

Using these approaches there has been considerable success in identifying DNA methylation changes related to the pathology of AD and other brain diseases[25,26,35]. To date, the largest EWAS in AD have been performed on bisuflite-treated DNA on the Illumina Infinium 450K Bead[25,26]. Although the two studies utilized different experimental designs, for example Lunnon *et* al performed a cross-tissue EWAS on ~120 individuals, whilst De Jager *et al* performed a single tissue EWAS on ~700 individuals, reassuringly both studies identified a number of overlapping loci[36]. Through recent advances in epigenomic studies, it is now possible to separately measure DNA methylation and DNA hydroxymethylation in the same samples by undertaking oxidative bisulfite treatment. Moreover, the recent release of the Illumina Infinium EPIC BeadChip array, means that nearly twice as many CpG sites can be investigated. The completion of the first EWAS of DNA hydroxymethylation and utilizing the Illumina Infinium EPIC BeadChip array are highly anticipated and are hoped to produce novel insights into the role of epigenetic mechanisms in AD pathology.

6. References

1. Seshadri, S. *et al.* Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* **303**, 1832-40 (2010).
2. Flores, K. *et al.* Genome-wide association between DNA methylation and alternative splicing in an invertebrate. *BMC Genomics* **13**, 480 (2012).
3. Wijsman, E.M. *et al.* Genome-wide association of familial late-onset Alzheimer's disease replicates BIN1 and CLU and nominates CUGBP2 in interaction with APOE. *PLoS Genet* **7**, e1001308 (2011).
4. Edwards, T.L. *et al.* Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *Ann Hum Genet* **74**, 97-109 (2010).
5. Abraham, R. *et al.* A genome-wide association study for late-onset Alzheimer's disease using DNA pooling. *BMC Med Genomics* **1**, 44 (2008).
6. Lambert, J.C. *et al.* Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet* **41**, 1094-9 (2009).
7. Harold, D. *et al.* Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet* **41**, 1088-93 (2009).
8. Kamboh, M.I. *et al.* Genome-wide association study of Alzheimer's disease. *Transl Psychiatry* **2**, e117 (2012).
9. Levinson, D.F. *et al.* Genome-wide association study of multiplex schizophrenia pedigrees. *Am J Psychiatry* **169**, 963-73 (2012).
10. Sherva, R. *et al.* Genome-wide association study of the rate of cognitive decline in Alzheimer's disease. *Alzheimers Dement* **10**, 45-52 (2014).
11. Moonesinghe, R., Khoury, M.J., Liu, T. & Ioannidis, J.P.A. Required sample size and nonreplicability thresholds for heterogeneous genetic associations. *Proceedings of the National Academy of Sciences* **105**, 617-622 (2008).
12. Evangelou, E. & Ioannidis, J.P.A. Meta-analysis methods for genome-wide association studies and beyond. *Nat Rev Genet* **14**, 379-389 (2013).
13. Ioannidis, J.P.A., Trikalinos, T.A. & Khoury, M.J. Implications of Small Effect Sizes of Individual Genetic Variants on the Design and Interpretation of Genetic Association Studies of Complex Diseases. *American Journal of Epidemiology* **164**, 609-614 (2006).
14. Henikoff, S. & Matzke, M.A. Exploring and explaining epigenetic effects. *Trends in Genetics* **13**(1997).
15. Lunnon, K. & Mill, J. Epigenetic studies in Alzheimer's disease: current findings, caveats, and considerations for future studies. *Am J Med Genet B Neuropsychiatr Genet* **162B**, 789-99 (2013).
16. Song, C.X., Yi, C. & He, C. Mapping recently identified nucleotide variants in the genome and transcriptome. *Nat Biotechnol* **30**, 1107-16 (2012).
17. Lunnon, K. *et al.* Variation in 5-hydroxymethylcytosine across human cortex and cerebellum. *Genome Biol* **17**, 27 (2016).
18. Schubeler, D. Function and information content of DNA methylation. *Nature* **517**, 321-6 (2015).
19. Hannon, E., Lunnon, K., Schalkwyk, L. & Mill, J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics* **10**, 1024-1032 (2015).
20. Braak, H. & Braak, E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol* **82**, 239-59 (1991).
21. Barrachina, M. & Ferrer, I. DNA methylation of Alzheimer disease and tauopathy-related genes in postmortem brain. *J Neuropathol Exp Neurol* **68**, 880-91 (2009).

22. Condliffe, D. *et al.* Cross-region reduction in 5-hydroxymethylcytosine in Alzheimer's disease brain. *Neurobiology of Aging* **35**, 1850-1854 (2014).
23. Chouliaras, L. *et al.* Consistent decrease in global DNA methylation and hydroxymethylation in the hippocampus of Alzheimer's disease patients. *Neurobioly of Aging* **34**, 2091-9 (2013).
24. Coppieters, N. *et al.* Global changes in DNA methylation and hydroxymethylation in Alzheimer's disease human brain. *Neurobiol Aging* **35**, 1334-44 (2014).
25. De Jager, P.L. *et al.* Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat Neurosci* **Sep;17**, 1156-63 (2014).
26. Lunnon, K. *et al.* Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nat Neurosci.* **17**(2014).
27. Sanchez-Mut, J.V. *et al.* Human DNA methylomes of neurodegenerative diseases show common epigenomic patterns. *Transl Psychiatry* **6**, e718 (2016).
28. Sanchez-Mut, J.V. *et al.* Promoter Hypermethylation of the Phosphatase DUSP22 Mediates PKA-Dependent TAU Phosphorylation and CREB Activation in Alzheimer's Disease. *Hippocampus* **24**, 363-368 (2014).
29. Watson, C.T. *et al.* Genome-wide DNA methylation profiling in the superior temporal gyrus reveals epigenetic signatures associated with Alzheimer's disease. *Genome Medicine* **8**, 5 (2016).
30. Bakulski, K.M. *et al.* Genome-wide DNA methylation differences between late-onset Alzheimer's disease and cognitively normal controls in human frontal cortex. *J Alzheimers Dis* **29**, 571-88 (2012).
31. Sambrook, J. & Russell, D.W. Purification of Nucleic Acids by Extraction with Phenol:Chloroform. *Cold Spring Harbor Protocols* **2006**, pdb.prot4455 (2006).
32. Frommer, M. *et al.* A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A* **89**, 1827-31 (1992).
33. Booth, M.J. *et al.* Quantitative Sequencing of 5-Methylcytosine and 5-Hydroxymethylcytosine at Single-Base Resolution. *Science* **336**, 934-937 (2012).
34. Pidsley, R. *et al.* Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol* **17**, 208 (2016).
35. Hannon, E. *et al.* An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation. *Genome Biology* **17**, 176 (2016).
36. Smith, A.R., Mill, J., Smith, R.G. & Lunnon, K. Elucidating novel dysfunctional pathways in Alzheimer's disease by integrating loci identified in genetic and epigenetic studies. *Neuroepigenetics* **6**, 32-50 (2016).

Table 1.1 Lysis Buffer Reagents

Name	Storage	pH	Final Concentration	Warnings
NaCl	Room temp		75mM	Irritant
Tris-HCl	Room temp	pH8	10mM	Corrosive Toxic
EDTA	Room temp	pH8	25mM	Irritant
10% SDS	Room temp		$1/20^{th}$ final volume	Irritant Flammable
H_2O	Room temp			Not Hazardous

Table 1.2 1x Te Solution (1L) – stable at room temperature

Name	Storage	pH	Volume (mL)	Warnings
1M tris-HCL	Room temp	pH8	10	Corrosive Toxic
0.5M EDTA	Room temp	pH8	0.2	Irritant
H_2O	Room temp		990	Not Hazardous

Filter with 0.5 micron filter and then autoclave before use.

Table 1.3 Other reagents needed for Phenol-Chloroform DNA extraction

Name	Storage	Concentration	Warnings
Proteinase K solution	-20°C	20mg/mL	Not Hazardous
Ribonuclease A (from bovine pancreas)	-20°C	20µg/mL	Not Hazardous
Phenol / Chloroform / Isoamyl Alcohol (PCI)	4-8°C	100%	Toxic Flammable Irritant
Chloroform ($CHCl_3$)	Room temp	100%	Irritant Toxic Flammable
Ethanol	-20°C	100%	Flammable
NaCl	Room temp	75mM	Irritant
Ethanol	Room temp	70%	Flammable
Elution Buffer (H_2O / Te)	Room temp	200-300µL per sample	Not Hazardous

Table 2.1 CT Conversion Reagent

Name	Volume (mL)	Supplier
M-Dissolving Buffer	0.5	Zymo
M-Dilution Buffer	3	Zymo
CT Conversion Reagent	Bottle	Zymo
Ultra-pure H_2O	9	User

Table 2.2 M-Wash Buffer

Name	Volume (mL)	Supplier
M-Wash Buffer	36	Zymo
100% Ethanol	144	User

Table 3.1 Magnetic Bead Binding Solution 1

Name	Volume (mL)	Supplier
Magnetic Bead Solution	0.24	CEGX
Binding Buffer 1	12	CEGX

Table 3.2 80% Acetonitrile (for 12 samples)

Name	Volume (mL)	Supplier
100% Acetonitrile	32	User
ultra-pure H_2O	8	CEGX

Table 3.3 Oxidation reaction

Component	Volume per reaction (µL)	
	oxBS	BS
Denatured DNA from step 12	**24**	**24**
Oxidant Solution	**1**	**-**
Ultra-Pure Water	**-**	**1**
Total volume	**25**	**25**

Table 3.4 Bisulfite Conversion Reaction Mix

Component	Volume per reaction (µL)
DNA solution (step 16)	**25**
Bisulfite Reagent Solution	**170**
Bisulfite Additive	**5**
Total volume	**200**

Table 3.5 Bisulfite conversion thermal profile

Step	Time	Temperature
Denaturation	5 minutes	95 °C
Incubation	20 minutes	60 °C
Denaturation	5 minutes	95 °C
Incubation	40 minutes	60 °C
Denaturation	5 minutes	95 °C
Incubation	165 minutes	60 °C
Denaturation	5 minutes	95 °C
Incubation	20 minutes	60 °C
Denaturation	5 minutes	95 °C
Incubation	40 minutes	60 °C
Denaturation	5 minutes	95 °C
Incubation	165 minutes	60 °C
Hold	Indefinite	20 °C

Table 3.6 Magnetic Bead Binding Solution 2

Name	Volume (mL)	Supplier
Magnetic Bead Solution	2.64	CEGX
Binding Buffer 2	220	CEGX

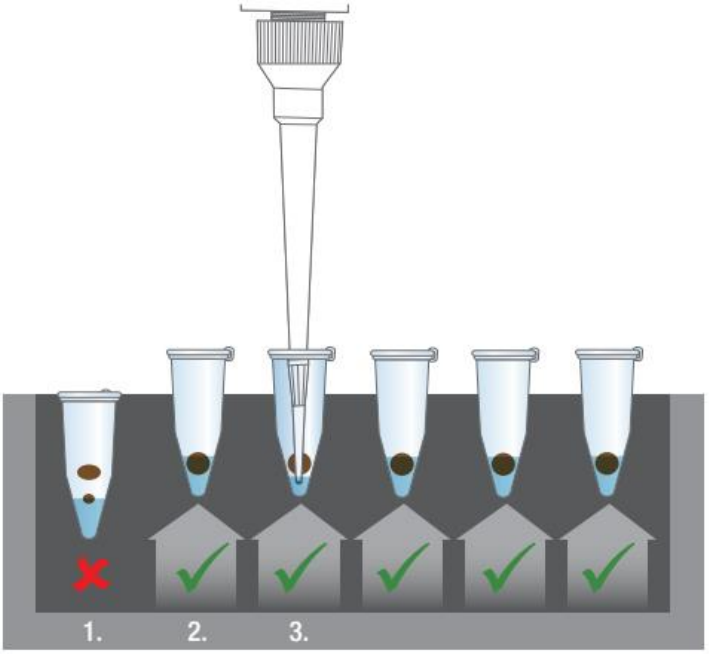
Table 3.7 Desulfonation Buffer

Name	Volume (mL)	Supplier
Desulfonation Buffer	Bottle	CEGX
100% Ethanol	150	User

Table 3.8 70% Ethanol (for 12 samples or 24 bisulfite conversions)

Name	Volume (mL)	Supplier
100% Ethanol	70	User
ultra-pure H2O	30	CEGX

Figure 1



APPENDIX 3 – PARALLEL PROFILLING OF DNA METHYLATION AND HYDROXYMETHYLATION HIGHLIGHTS MEUROPATHOLOGY-ASSOCIATED VARIATION IN ANK1 AND THIRTEEN NOVEL LOCI IN ALZHEIMER'S DISEASE ENTORHINAL CORTEX

Parallel profiling of DNA methylation and hydroxymethylation highlights neuropathology-associated epigenetic variation within *ANK1* and thirteen novel loci in Alzheimer's disease entorhinal cortex.

Adam R Smith[1], Rebecca G Smith[1], Eilis Hannon[1], Janou A Y Roubroeks[1,2], Joe Burrage[1], Claire Troakes[4], Safa Al-Sarraj[4], Jonathan Mill[1], Daniel L van den Hove[2,3], Katie Lunnon[1,*]

[1] University of Exeter Medical School, Exeter University, Exeter, UK.

[2] School for Mental Health and Neuroscience, Department of Psychiatry and Neuropsychology, Maastricht University Medical Centre, Maastricht, The Netherlands

[3] Division of Molecular Psychiatry, Laboratory of Translational Neuroscience, Department of Psychiatry, Psychosomatics and Psychotherapy, University of Würzburg, Würzburg, Germany

[4] Institute of Psychiatry, King's College London, London, UK.

* Corresponding author: Katie Lunnon, University of Exeter Medical School, RILD Building Level 4, Royal Devon and Exeter Hospital, Barrack Rd, Exeter. EX2 5DW. UK. E-mail: k.lunnon@exeter.ac.uk

ABSTRACT

Background

Alzheimer's disease is a progressive neurodegenerative disorder that is hypothesized to involve epigenetic dysfunction. Previous epigenome-wide association studies in Alzheimer's disease have been unable to distinguish between two DNA modifications: DNA methylation and hydroxymethylation. DNA hydroxymethylation has been shown to be enriched in the human brain, although its role in Alzheimer's disease has not yet been fully explored. Here we utilize oxidative bisulfite conversion, in conjunction with the Illumina 450K BeadChip, to identify neuropathology-associated differential DNA methylation and hydroxymethylation in two distinct anatomical regions of the brain (entorhinal cortex and cerebellum) dissected from post-mortem brain samples of donors with varying degrees of Alzheimer's disease pathology.

Results

We confirm altered DNA methylation levels in *ANK1*, and highlight 13 novel genes (*TMEM53, KCNJ9, H3F3A, OPN3, NHEJ1, ACVR2B, ROBO2, CGGBP1, RAPGEF6, GFRA1, FHOD1, FOXF1, KCNC3* and *APP*) not previously identified in Alzheimer's disease epigenome-wide association studies, which show significant interactions in DNA methylation and hydroxymethylation levels. We replicate significant *ANK1* disease-associated hypermethylation and hypohydroxymethylation patterns in an independent cohort using oxidative-bisulfite pyrosequencing.

Conclusions

Our study represents the first cross-tissue epigenome-wide association study of both DNA methylation and hydroxymethylation in Alzheimer's disease. We demonstrate that previous estimates of DNA hypermethylation in *ANK1* in Alzheimer's disease were underestimates as it is confounded by hypohydroxymethyation. Our study also identifies 13 novel genes that display DNA modification alterations in Alzheimer's disease, which would not have been identified using standard bisulfite based approaches.

KEYWORDS

Alzheimer's Disease (AD)

Brain

Cerebellum (CER)

DNA methylation (5mC)

DNA hydroxymethylation (5hmC)

Entorhinal Cortex (EC)

Epigenetics

Epigenome wide association study (EWAS)

Illumina Infinium 450K Beadarray

BACKGROUND

Alzheimer's disease (AD) currently affects approximately 35 million people worldwide [1]. It is a progressive neurodegenerative disease that leads to neuronal cell loss and results in severe cognitive decline. The characteristic hallmarks of AD include neurofibrillary tangles (NFTs) of hyperphosphorylated tau and amyloid beta (Aβ) plaques [2]. Although Mendelian inheritance of mutations in the *APP*, *PSEN1* and *PSEN2* genes have been demonstrated in early-onset familial AD cases, these only account for ~5% of disease incidence [3]. The majority of AD cases are sporadic, occur late in life and have, as yet, no defined etiology, with common single nucleotide polymorphisms (SNPs) accounting for only a third of disease risk [4]. Epigenetic processes mediate the reversible regulation of gene expression, occurring independently of DNA sequence variation and orchestrate a diverse range of important neurobiological processes in the brain; with DNA methylation (5-methylcytosine - 5mC) being the best characterized and most stable epigenetic modification [5]. Two recent epigenome-wide association studies (EWAS) have utilized the Illumina 450K methylation array (450K array) to demonstrate robust and reproducible changes in DNA methylation at a number of loci in the AD brain [6, 7], including the *ANK1* gene. One caveat of this approach is the utilization of bisulfite-treated (BS) DNA, which converts unmodified cytosines to thymine (post PCR), whilst methylated cytosines remain unchanged. Although the focus of these studies has been on changes in DNA methylation in AD, BS-based approaches cannot distinguish this modification from DNA hydroxymethylation (5-hydroxymethylcytosine - 5hmC). 5hmC has been previously identified at high levels in the developing [8] and adult brain [9], particularly in neurons [10] and, as such, may represent an important epigenetic mark to profile in the context of neurodegenerative diseases. Furthermore, levels of 5hmC potentially mask the true abundance of 5mC at specific loci in the genome, confounding existing EWAS analyses of AD.

Recent studies have demonstrated that oxidative BS (OxBS) treatment enables the detection of 5hmC as thymine, and therefore by running matched BS- and OxBS-treated samples in parallel we can generate a quantitative measurement for total DNA modifications (BS data), DNA methylation (OxBS data) and, by proxy, DNA hydroxymethylation (BS data – OxBS data) [11]. We have previously utilized this method in conjunction with 450K arrays to profile 5mC and 5hmC levels in parallel in post-mortem brain tissue from non-demented individuals to demonstrate region specific differences in DNA modifications [11]. Here we perform the first cross-tissue EWAS of DNA methylation and DNA hydroxymethylation independently in multiple tissues from 96 donors, representing the spectrum of AD pathology defined by Braak staging, a standardized measure of NFT burden determined at autopsy [12], ranging from no AD pathology (Braak 0) to late stage AD (Braak VI). From each donor, we analyzed the entorhinal Cortex (EC), a region of the brain affected early in the disease [2] and the cerebellum (CER), which is relatively spared from pathology.

RESULTS AND DISCUSSION

AD entorhinal cortex is characterized by global hypermethylation and hypohydroxymethylation

The majority of studies investigating DNA hydroxymethylation in AD have used immuno-histochemical based methods to quantify global levels [13-17]. One study reported a global decrease in 5hmC in the EC and CER in AD, but no change in 5mC [18]. However, other studies have reported global reductions in both modifications in the hippocampus [13] and another study highlighted increased levels of 5mC and 5hmC in the middle frontal gyrus and middle temporal gyrus [16]. Although the 450K array primarily interrogates CG rich regions, which are, in general, unmethylated, we were interested to see if there were any global changes across these regions. To investigate these global levels of DNA modifications in our data we averaged levels across all probes that passed quality control (QC) (*n* = 367,480),

identifying a global decrease in 5hmC (P = 0.040) and increase in 5mC (P = 0.039), with overall no net change in total modifications (P = 0.838) in the EC with respect to Braak stage. Conversely, no significant Braak stage associated alterations were seen for global levels of 5hmC (P = 0.659), 5mC (P = 0.916) or total modifications (P = 0.366) in the CER (P > 0.05).

Highly reproducible alterations in cytosine modifications are detectable in AD cortex

Initially, we focused on identifying differences in total DNA modification levels (BS data: **Supplementary Table 1**) in the EC that were associated with Braak stage, whilst controlling for age and sex above our nominated significance threshold ($P < 5 \times 10^{-5}$). To assess the reproducibility of our results, we first compared the effect size of these Braak stage-associated differentially modified positions (DMoPs) from the BS data (**Supplementary Table 1**), with BS data previously generated on an independent set of samples from the EC in AD (validation cohort 1) [7]. There was a significant excess of consistent direction of effect between the two studies (**Figure 1A**: sign test P =0.043), with the effect sizes highly correlated between studies (r = 0.71, $P = 4.85 \times 10^{-6}$). Similarly, there was also a significant excess of consistent direction of effect between the two studies when examining the effect size of previously published Braak-associated DMoPs ($P < 5 \times 10^{-5}$) in the EC [7] with the effect size of the same probes in the current study (**Figure 1B**: sign test $P = 9.77 \times 10^{-04}$), with the effect sizes highly correlated between studies (r = 0.78, $P = 3.50 \times 10^{-4}$). This suggests that although different BS conversion methods and donors were used, there are consistent and reproducible changes in total DNA modifications in the EC of donors with AD.

In contrast, when we compared the effect size of the Braak-associated DMoPs in the CER ($P < 5 \times 10^{-5}$) in the current study (**Supplementary Table 2**) to BS data previously published in the CER in AD [7] we found no consistent direction of effect (**Supplementary Figure 1A**: sign test P = 0.267). Similarly, the effect sizes of the previously published CER DMoPs (P <

5 x 10^{-5}) [7] showed no consistent direction of effect with the same probes in the current study (**Supplementary Figure 1B**: sign test *P* = 0.188). This suggests that regions of the brain that are relatively protected from neurodegeneration, such as the CER, do not display consistent disease-associated epigenetic changes in total DNA modifications in either a global or loci-specific manner.

Locus specific changes in DNA methylation and hydroxymethylation occur in AD

We were interested to identify specific loci that showed Braak-associated differential DNA methylation and hydroxymethylation in the EC (**Supplementary Tables 3-4**; **Figure 2A**). We identified three Braak-associated differentially methylated positions (DMPs) that reached experiment-wide significance ($P < 2.2 \times 10^{-7}$), with a further 72 loci reaching our nominal significance threshold of $P < 5 \times 10^{-5}$. Although no Braak-associated differentially hydroxymethylated positions (DHPs) reached experiment wide significance, we did identify 57 loci that reached our nominal significance threshold. Of the three experiment-wide significant DMPs, the first and third most significant were found on chromosome 3 and resided in *CGGBP1* and *GNL3* respectively, whilst the second most significant DMP was located in *WNT5B* on chromosome 12. *CCGBP1* is known to influence the expression of the *FMR1* gene, which is associated with the fragile X mental retardation syndrome [19], whereas *GNL3* is hypothesized to interact with p53, a tumor suppressor protein [20]. *WNT5B* is part of the wnt gene family and disruption of the wnt signaling pathway has been previously implicated in neurodegeneration [21, 22]. In the CER we only identified seven Braak-associated DMPs (**Supplementary Table 5**) and 12 DHPs (**Supplementary Table 6**) that reached our nominal significance threshold ($P < 5 \times 10^{-5}$).

Neuropathology-associated DMPs and DHPs in the EC feature in key biological pathways

We performed a pathway analysis of the genes annotated to DMPs and DHPs in the EC that reached our nominal significance threshold ($P < 5 \times 10^{-5}$) (**Supplementary Table 7** and **Supplementary Table 8**, respectively). Relevant pathways of interest included axon midline choice point recognition" and "axon choice point recognition", both of which were enriched for DMPs and DHPs (GO:0016199: 5mC $P = 2.30 \times 10^{-4}$, 5hmC $P = 1.30 \times 10^{-4}$; GO:0016198: 5mC $P = 4.20 \times 10^{-4}$, 5hmC $P = 2.55 \times 10^{-4}$), "regulation of synapse organization" (GO:0050807: 5hmC $P = 9.11 \times 10^{-4}$), "synapse organization" (GO:0050808: 5hmC $P = 0.001$) and "synapse assembly" (GO:0007416: 5hmC $P = 0.002$), which were all enriched for DHPs. This suggests that synaptic activity is a major pathway that displays altered levels of 5hmC in AD. Interestingly, previous work by Khare *et al.* has shown that 5-hmC is enriched in genes with synapse-related functions in both human and mouse brain [23].

Neuropathology-associated changes in 5mC and 5hmC are seen across gene features

Next, we compared the genic location of the most significant hyper- or hypo-(hydroxy)methylated neuropathology-associated loci in the EC ($P < 5 \times 10^{-5}$), to all neuropathology-associated DMPs or DHPs ($P < 0.05$), with relation to gene features that have been described by Slieker *et al* [24] (**Supplementary Figures 2-3; Supplementary Tables 9-10**). We demonstrated an enrichment for the most significant DMPs and DHPs across all gene features, most strikingly Braak-associated hypomethylation and hyperhydroxymethylation appears to be enriched in the CGIs of proximal promotors ($P < 1 \times 10^{-330}$). This shows that there could be a change from DNA methylation to DNA hydroxymethylation in these regions in AD, despite a global decrease (across regions covered by the array) in 5hmC with disease.

A number of loci show display significant negative correlations in neuropathology-associated DNA methylation and hydroxymethylation changes

Next, we sought to explore whether there was any overlap in disease-associated DMoPs, DMPs and DHPs. We compiled a list of loci that showed a disease-associated alteration in more than one analysis ($P < 5 \times 10^{-5}$) (**Figure 2B**; **Table 1**). We identified 14 CpG sites (annotated to 14 genes) that showed significant changes across more than one analysis (total modifications, 5mC alone and 5hmC alone) (**Supplementary Figure 4**). Of the 14 CpGs nominated, 13 showed no nominally significant ($P > 0.05$) difference in total DNA modifications. This is consistent with findings from previous EWAS using BS DNA, which had also not shown these loci to be associated with AD-neuropathology. Although we observed no difference in total DNA modifications in disease for these 13 loci, we did observe neuropathology-associated differences in both 5mC and 5hmC, with in each instance the levels of both modifications being highly negatively correlated with each other (**Supplementary Figure 5**) and a significant interaction between the two modifications. Of particular interest are differences at a site annotated to the *APP* gene (cg12144476), given that mutations in this gene are associated with early-onset, familial AD. In the current study we saw a decrease in 5mC in *APP* with Braak stage (**Supplementary Figure 4N**; Effect size (difference (Δ) Braak 0 to Braak VI) = -1.39, $P = 4.85 \times 10^{-06}$), with a concurrent parallel increase in 5hmC (**Supplementary Figure 4N**; $\Delta = 1.44$, $P = 1.61 \times 10^{-05}$) demonstrating a significant interaction between the two modifications (**Supplementary Figure 5M**; $P = 1.77 \times 10^{-10}$) and no net change in total DNA modifications.

ANK1 shows differential DNA methylation in the AD EC

Next, we performed an analysis to identify spatially correlated regions consisting of >2 neighboring DMPs (differentially methylated regions - DMRs) or DHPs (differentially hydroxymethylated regions – DHRs) within a 500bp sliding window (*comb-p* [25]). We did not identify any DHRs; however, given that DNA hydroxymethylation is known to be enriched in areas of low CG density [11], this is perhaps not surprising. We identified four DMRs that reached a Sidak-corrected *P* value < 0.05 (**Table 2**). This included an eight probe DMR

within the *WNT5A* gene (Sidak-corrected $P = 3.17 \times 10^{-4}$), a three probe DMR within the *TRAF3* gene (Sidak-corrected $P = 0.022$), a two probe DMR within the *ANK1* gene (Sidak-corrected $P = 0.025$) and a six probe DMR within the *ARID5B* gene (Sidak-corrected $P = 0.029$). The two DMPs we identified in *ANK1* in our regional analysis have previously been nominated as DMoPs in AD EC in two independent EWAS using BS-treated DNA [6, 7]. In the current study, we again identified a significant increase in total DNA modifications in AD at both these two sites in the EC (**Figure 2C**: cg11823178: $\Delta = 6.69$, $P = 4.68 \times 10^{-4}$ and **Figure 2D**: cg05066959: $\Delta = 10.98$, $P = 8.37 \times 10^{-6}$). However, this is the first study to demonstrate that these are specifically attributable to increased 5mC (cg11823178: $\Delta = 9.31$, $P = 2.61 \times 10^{-5}$; cg05066959: $\Delta = 13.71$, $P = 8.03 \times 10^{-5}$), with a small, but non-significant decrease in 5hmC at both sites (cg11823178: $\Delta = -3.20$, $P = 0.068$; cg05066959: $\Delta = -2.72$, $P = 0.349$). This suggests that previous studies of DNA methylation in *ANK1* have underestimated increments in 5mC in disease due to confounding by 5hmC.

Oxidative-bisulfite pyrosequencing validation of ANK1

To further confirm that *ANK1* epigenetic changes associated with neuropathology involve DNA hypermethylation, and not hyperhydroxymethylation, we used OxBS pyrosequencing to quantify DNA modifications across an extended region spanning eight CpG sites, including cg11823178 and cg05066959. DNA used for this study was from an independent collection of EC tissue ($n = 96$) obtained from the Thomas Willis Oxford Brain Collection (validation cohort 2) (**Supplementary Table 11**) [26]. In this cohort we did not have access to Braak stage III or IV samples, and as such analyzed data using a case (Braak V-VI) /control (Braak)-II) analysis model. Of the eight CpG sites assessed, seven where characterized by significant ($P < 0.05$) AD-associated hypermethylation (**Figure 3A**). Interestingly, significant AD-associated hypohydroxymethylation was also seen at four of the eight CpG sites (**Figure 3B**), with total DNA modifications showing a remarkably similar profile to previously published studies [7] (**Figure 3C**). When we averaged total DNA modifications, 5mC and

5hmC levels across the eight CpG sites in the 120bp region, we saw a significant increase in total modifications (Δ = 0.79, P = 1.12 x 10^{-4}) and 5mC (Δ = 0.87, P = 1.23 x 10^{-05}) and a trend towards a decrease in 5hmC, (Δ = -0.40, P = 0.058) in individuals with AD, compared to controls (**Figure 3D**). This data suggests that previous studies of *ANK1* DNA methylation in AD have underestimated the disease-associated changes in 5mC.

Several recent studies have reported the links between epigenetic alterations at the *ANK1* locus and the development of AD [6, 7, 27]. The ANK1 protein is a plasma-bound membrane protein that contains an ankyrin repeat domain, which modulates interactions between cytoskeletal and membrane proteins [28]. One of the main functions of *ANK1* is compartmentalization and maintenance of the plasma membrane and it is possible that the altered expression of this gene could lead to cell membrane dysfunction in AD [7]. It has been shown that microglia, the brain's resident macrophages, have >4-fold increase in *ANK1* expression in AD compared to control brain samples [27], suggesting it is possible that *ANK1* DNA methylation levels will have a downstream effect on immune and inflammatory regulation in the brain.

CONCLUSIONS

Our study represents, to our knowledge, the first large scale cross-tissue study of 5mC and 5hmC, at base-pair level resolution in AD, using matched DNA from both affected and unaffected brain regions, utilizing independent study cohorts and two independent technologies (OxBS-Illumina 450K array and OxBS-pyrosequencing). Two previous studies of DNA hydroxymethylation in AD have been recently published. The first used a selective chemical labelling technique to enrich for 5hmC, and then sequenced the captured libraries from a small cohort of 30 individuals with either AD, mild cognitive impairment (MCI) or no dementia [29]. The study highlighted a number of AD-associated DHPs and demonstrated that gene body DNA hydroxymethylation was positively correlated with cis-acting gene

expression. One limitation of that study was that the analysis was unable to discriminate between 5hmC and 5mC because of the low sequencing resolution. The second study, by Ellison and colleagues, utilized reduced representation hydroxymethylation profiling (RRHP) to analyse 5hmC levels in AD hippocampus [30]. Although this study assessed DNA hydroxymethylation at >2million sites, the study was limited to just three AD cases and two age-matched control subjects. In our study the overlap between DMoPs identified across our sample cohorts and those identified by Lunnon *et al.* [7], despite the use of independent study cohorts, suggests that our study was adequately powered to detect robust AD-associated differences that can be replicated in other studies. Our analyses from multiple independent cohorts provide further evidence for a role for DNA hypermethylation, coupled with hypohydroxymethylation, across a region in *ANK1* in AD-associated neuropathology. It indicates that previous estimates of hypermethylation in *ANK1* were underestimates as it is potentially confounded by this hypohydroxymethyation. Our study also nominates 13 novel loci that show no change in overall DNA modifications, but display significant (opposing) differences in 5mC and 5hmC in AD. As these genes were not identified in standard BS-treated DNA, these findings highlight the utility of our method for identifying novel epigenetically dysregulated genes in AD. Although some of the AD-associated loci identified in our study do not have a mean 5hmC level across all samples >0, this likely reflects experimental noise resulting from the calculation of 5hmC from two separate arrays. However, given that the loci we identified resided within genes previously implicated in AD, such as *APP*, it is likely that our data is highlighting novel epigenetically altered loci in disease. To conclude, these novel genes warrant further research in AD.

METHODS

Subjects and samples

For BS and OxBS Illumina 450K profiling (discovery cohort) we used matched brain samples collected from the EC and CER from 96 individuals archived in the MRC London Neurodegenerative Disease Brain Bank (http://www.kcl.ac.uk/iop/depts/cn/research/MRC-London-Neurodegenerative-Diseases-Brain-Bank/MRC-London-Neurodegenerative-Diseases-Brain-Bank.aspx). These samples were not utilized in previously published AD EWAS publications [7, 31, 32]. Our first validation cohort consisted of previously published [7] BS EWAS data generated in matched EC (*n* = 104) and CER (*n* = 108) brain tissue. Our second validation cohort (for pyrosequencing) consisted of matched EC and CER samples from 96 individuals archived in the Thomas Willis Oxford Brain Collection (http://www.medsci.ox.ac.uk/optima/information-for-patients-and-the-public/the-thomas-willis-oxford-brain-collection) [26]. For all cohorts, individuals had varying degrees of AD pathology (Braak Stage 0-VI) and all AD patients were over the age of 65 at clinical diagnosis. All samples were dissected by trained specialists, snap-frozen and stored at -80°C. Further information about all samples is provided in **Supplementary Table 11**. For the discovery cohort and validation cohort 2, genomic DNA was isolated from ~100mg of each dissected brain region using a standard phenol-chloroform extraction method, and tested for degradation and purity prior to analysis.

Methylomic and hydroxymethylomic profiling

1µg DNA from each sample (discovery cohort and validation cohort 2) was treated with sodium BS and OxBS in parallel using the true-methyl CEGX 96 kit (CEGX, Cambridge, UK) according to the manufacturer's standard protocol. Briefly, DNA samples were split, with half being oxidized (OxBS) and the remainder (BS) going through a mock oxidization step, before all being BS treated. All samples were then processed using the Illumina Infinium HumanMethylation450K BeadChip (450K array) (Illumina Inc, CA, USA) according to the manufacturer's instructions, with minor amendments and quantified using an Illumina HiScan System (Illumina, CA, USA). Corresponding OxBS and BS treated DNA for the same sample were run together on the same BeadChip. Samples were assigned a unique code for the

purpose of the experiment and grouped by tissue. Samples were randomized in their OxBS and BS pairs with respect to sex and disease status to avoid batch effects, and processed in batches of 12 BeadChips. Illumina Genome Studio software was used to extract the raw signal intensities of each probe (without background correction or normalization).

Data Analysis

All computations and statistical analyses were performed using R 3.3.2 [33] and Bioconductor 3.5 [34]. Signal intensities were imported into R using the methylumi package [35] as a methylumi object. Initial QC checks were performed using functions in the methylumi package to assess concordance between reported and genotyped gender. Non-CpG SNP probes on the array were also used to confirm that both brain regions were sourced from the same individual where expected. Probes with common (minor allele frequency (MAF) > 5%) SNPs in the CG or single base extension position or probes that are nonspecific or mismapped were flagged and discarded from our results [36]. Data was pre-processed in the R package *wateRmelon* using the *dasen* function as previously described [37]. The pfilter function was used to remove samples where 5% of sites with a detection $P >$ 0.05. Specific sites were also removed if the beadcount was less than three in 5% of samples or if 1% of the samples had a detection $P > 0.05$ at that position. After these QC steps 367,480 probes were taken forward for analysis. Array data for each tissue was normalized separately and the analysis was performed separately by tissue. To calculate 5hmC levels in a sample, the OxBS signal was subtracted from the BS signal. As we have previously reported a small proportion of probes in each sample were characterized by a negative BS-oxBS value, likely resulting from technical variance inherent in the Illumina array protocol [11]. Therefore probes with a mean negative 5hmC value across all samples were included in our analysis, however we annotated tables to denote that the mean 5hmC level was >0 (5hmC Positive probe column). The effects of age and sex were regressed out before subsequent analysis. For the identification of DMoPs, DMPs and DHPs specifically

altered with respect to neuropathological measures of AD, we performed a quantitative analysis where samples were analyzed with respect to Braak stage using a linear regression model. To ensure that the top DMoPs, DMPs and DHPs were not a result of cell proportion differences, the *CETS* package in R [38] was used to determine neuron/glia proportions and tables annotated with an additional column including this value as an additional co-variate (*P* CETS corrected). Data was analyzed separately for each brain region using linear regression and probes were ranked by *P* value, and Q-Q plots assessed to check for *P* value inflation (see **Supplementary Figure 6**).

Identifying enrichment in specific gene features

To investigate whether our top neuropathology-associated DMPs were enriched at specific gene features, we used a Fisher's exact test to test for an enrichment of either significant hypermethylated or hypomethylated loci ($P < 5 \times 10^{-5}$) compared to all neuropathology-associated DMPs ($P < 0.05$) at each gene feature as described by Slieker et al [24]. Similarly to investigate whether our top neuropathology-associated DHPs were enriched at specific gene features, we again used a Fisher's exact test to test for an enrichment of either significant hyperhydroxymethylated or hypohydroxymethylated loci ($P < 5 \times 10^{-5}$) compared to all neuropathology-associated DHPs ($P < 0.05$) at each Slieker gene feature.

Differentially methylated regions

To identify DMRs and DHRs we identified spatially correlated p-values within our data using the Python module *comb-p* [25] to group spatially correlated CpGs within a 500bp sliding window with a significance threshold of $P = 0.01$.

Pathway analysis

Pathway analysis of the genes reaching our significance threshold ($P < 5 \times 10^{-5}$) was performed using the missmethyl 1.10.0 [39] package from Bioconductor. Pathways were included if they met the following criteria: (1) the number of genes in the gene list was greater than one and less than 2000, (2) more than one gene from our analysis was present in the pathway and (3) the *P* value for the association was smaller than 0.05.

Targeted Replication using Bisulfite Pyrosequencing

Bisulfite pyrosequencing was used to quantify DNA methylation across eight individual *ANK1* CpG sites, including cg05066959 and cg11823178, spanning from 41519302 to 41519420 within chromosome 8 (hg19). A single amplicon (246bp) was amplified using primers designed using the PyroMark Assay Design software 2.0 (Qiagen, UK) as previously described [7], and sequenced using two sequencing primers to maximize coverage across eight CpG sites within a 120bp region. DNA methylation was quantified in validation cohort 2 using the Pyromark Q24 system (Qiagen, UK) following the manufacturer's standard instructions and the Pyro Q24 CpG 2.0.6 software. Data was adjusted for the effects of age and sex. For this sample cohort we had no samples with middle stage AD (Braak III-IV). In this instance rather than performing a linear regression analysis across Braak stages we instead compared control samples (Braak scores 0-II: $n = 48$) to AD samples (Braak scores V-VI: $n = 48$) for total DNA modifications, 5mC and 5hmC at eight individual CpG sites, and across the 120bp amplicon.

DECLARATIONS

Ethics approval and consent to participate

Ethical approval for the study was provided by the NHS South East London REC 3.

Consent for Publication

Not applicable.

Availability of data and materials

The datasets generated and analyzed during the current study are available from GEO under accession number GSE105109.

Competing Interests

The authors declare that they have no competing interests.

Funding

This work was funded by an Alzheimer's Association US New Investigator Research Grant (NIRG-14-320878) to KL, a PhD studentship from BRACE (Bristol Research into Alzheimer's and Care of the Elderly) to KL, a major project grant from the Alzheimer's Society UK (AS-PG-14-038) to KL and project grant funds from the Medical Research Council (MRC) (MR/N027973/1) to KL as part of the Joint Programme—Neurodegenerative Disease Research (JPND) funded consortium project "Targeting epigenetic dysregulation in the brainstem in Alzheimer's Disease (EPI-AD)" to DLvdH from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 643417. More generally the EPI-AD project is supported through the following funding organizations under the aegis of JPND—http://www.jpnd.eu: The Netherlands Organisation for Health Research and Development (ZonMw) [The Netherlands], Medical Research Council [United Kingdom], German Federal ministry of Education and Research (BMBF) [Germany], National Research Fund (FNR) [Luxembourg].

Author's Contributions

AS and JB conducted laboratory experiments. AS, RS, EH and JR undertook data analysis and bioinformatics. CT and SAS provided samples for analysis. KL, DLvdH and JM conceived and supervised the project. AS and KL drafted the manuscript. All authors read and approved the final submission

Acknowledgements

We thank the Brains for Dementia Research Initiative (Alzheimer Brain Bank UK) and the donors and families who made this research possible.

LIST OF ABBREVIATIONS

450K array	Illumina 450K methylation array
5hmC	5-hydroxymethylctosine
5mC	5-methylcytosine
Aβ	Amyloid beta
AD	Alzheimer's disease
BS	Bisulfite
CER	Cerebellum
CGI	CpG island
DHP	Differentially hydroxymethylated position
DHR	Differentially hydroxymethylated region
DMR	Differentially methylated region
DMoP	Differentially modified position
DMP	Differentially methylated position
EC	Entorhinal cortex
EWAS	Epigenome-wide associations tudy
MAF	Minor allele frequency
NFT	Neurofibrillary tangle
OxBS	Oxidative bisulfite
QC	Quality control
SNP	Single nucleotide polymorphism

FIGURE LEGENDS

Figure 1: Braak-associated EC DMoPs identified in this study are consistent with those identified in previous analyses of AD brain. (**A**) There was a consistent direction of effect for Braak-associated DMoPs ($P < 5 \times 10^{-5}$) identified in the EC in the discovery cohort when compared with the same probes in a previously published dataset of Braak-associated DMoPs in the EC (validation cohort 1) (sign test P =0.043), with the effect sizes highly correlated between studies (r = 0.71, P = 4.85×10^{-6}). (**B**) Similarly, there was a consistent direction of effect for previously published Braak-associated DMoPs in the EC ($P < 5 \times 10^{-5}$) (validation cohort 1) when compared with the same probes in the current study (sign test $P = 9.77 \times 10^{-04}$), with the effect sizes highly correlated between studies (r = 0.78, $P = 3.50 \times 10^{-4}$).

Figure 2: **Alterations in total DNA modifications, DNA methylation and DNA hydroxymethylation are associated with AD neuropathology in the EC.** (**A**) A Manhattan plot of association between DNA methylation (purple, above X axis) and DNA hydroxymethylation (yellow, below X axis) with Braak stage in the EC highlights associations at loci across the genome. We identified three DMPs reaching experiment wide significance (red line; $P < 2.2 \times 10^{-7}$), and a further 72 DMPs reaching our nominal significance threshold (green line; $P < 5 \times 10^{-5}$). Although no DHPs reached experiment wide significance, we identifed 57 DHPs at our nominal significance threshold. (**B**) A circos-plot highlighting Braak-associated DMoPs (grey– inner circle), DMPs (purple– middle circle) and DHPs (yellow – outer circle). Loci that showed a disease-associated epigenetic alteration in more than one analysis ($P < 5 \times 10^{-5}$) are inscribed around the outside. (**C-D**) Scatter plot of beta values, corrected for age and gender, against Braak stage for probes identified in *ANK1*. Purple circles represent 5mC levels (OxBS), yellow circles represent 5hmC levels (BS-OxBS).

Regression lines are shown for 5mC (purple line), 5hmC (yellow line) and total DNA modifications (grey line). Solid regression lines indicate P <0.05, whilst dashed lines indicate P > 0.05. Shown are (**C**) cg11823178 (5mC: r = 0.440, P = 2.61 x 10^{-5}; 5hmC: r = -0.199, P = 0.068; total modifications: r = 0.356, P = 4.68 x 10^{-4}) and (**D**) cg05066959 (5mC: r = 0.415, P = 8.03 x 10^{-5}; 5hmC: r = -0.103, P = 0.349; total modifications: r = 0.440, P = 8.37 x 10^{-6}).

Figure 3: Sites in *ANK1* are characterized by significant DNA hypermethylation and hypohydroxymethylation in AD across an extended region. Using BS and OxBS pyrosequencing we determined (**A**) 5mC (**B**) 5hmC and (**C**) total DNA modification levels in the entoECrhinal cortex in AD samples (Braak V-VI) compared to control samples (Braak 0-II) in validation cohort 2. We assayed a 120bp region containing cg05066959 (chr8:41519308) and cg1182378 (chr8:41519399). We demonstrated significant (P < 0.05) neuropathology-associated hypermethylation at seven of the eight CpG sites, hypohydroxymethylation at four of the eight CpG sites, and increased total DNA modifications at four of the eight sites. (**D**) Global analysis of all sites within the 120bp amplicon region highlighted a significant increase in total modifications (P = 1.12 x 10^{-4}) and 5mC (P = 1.23 x 10^{-05}) and a trend towards a decrease in 5hmC (P = 0.058) in individuals with AD, compared to controls. Data is represented as mean (±SEM) Key: *P < 0.05, **P < 0.01, ***P < 0.005.

REFERENCES

1. Martin Prince, et al., *World Alzheimer Report 2015 - The Global Impact of Dementia, an analysis of prevalence,incidence,cost and trends.* Alzheimer's Disease International, 2015: p. 1-87.
2. Serrano-Pozo, A., et al., *Neuropathological Alterations in Alzheimer Disease.* Cold Spring Harbor Perspectives in Medicine:, 2011. **1**(1): p. a006189.
3. Association, A.s., *Alzheimer's Disease Facts and Figures 2016.* Alzheimer's & Dementia, 2016. **12**(4): p. 1-84.
4. Lambert, J.C., et al., *Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease.* Nat Genet, 2013. **45**(12): p. 1452-8.
5. Lunnon, K. and J. Mill, *Epigenetic studies in Alzheimer's disease: current findings, caveats, and considerations for future studies.* Am J Med Genet B Neuropsychiatr Genet, 2013. **162B**(8): p. 789-99.
6. De Jager, P.L., et al., *Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci.* Nat Neurosci, 2014. **Sep;17**(9): p. 1156-63.
7. Lunnon, K., et al., *Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease.* Nat Neurosci, 2014. **Sept; 17**(9): p. 1164-70.
8. Spiers, H., et al., *5-hydroxymethylcytosine is highly dynamic across human fetal brain development.* BMC Genomics, 2017. **18**(1): p. 738.
9. Munzel, M., et al., *Quantification of the sixth DNA base hydroxymethylcytosine in the brain.* Angew Chem Int Ed Engl, 2010. **49**(31): p. 5375-7.
10. Kriaucionis, S. and N. Heintz, *The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain.* Science, 2009. **324**(5929): p. 929-30.
11. Lunnon, K., et al., *Variation in 5-hydroxymethylcytosine across human cortex and cerebellum.* Genome Biol, 2016. **17**: p. 27.
12. Braak, H. and E. Braak, *Neuropathological stageing of Alzheimer-related changes.* Acta Neuropathol, 1991. **82**(4): p. 239-59.
13. Chouliaras, L., et al., *Consistent decrease in global DNA methylation and hydroxymethylation in the hippocampus of Alzheimer's disease patients.* Neurobioly of Aging, 2013. **34**(9): p. 2091-9.
14. Condliffe, D., et al., *Cross-region reduction in 5-hydroxymethylcytosine in Alzheimer's disease brain.* Neurobiology of Aging, 2014. **35**(8): p. 1850-1854.
15. Lashley, T., et al., *Alterations in global DNA methylation and hydroxymethylation are not detected in Alzheimer's disease.* Neuropathol Appl Neurobiol, 2015. **41**(4): p. 497-506.
16. Coppieters, N., et al., *Global changes in DNA methylation and hydroxymethylation in Alzheimer's disease human brain.* Neurobiol Aging, 2014. **35**(6): p. 1334-44.
17. Celarain, N., et al., *TREM2 upregulation correlates with 5-hydroxymethycytosine enrichment in Alzheimer's disease hippocampus.* Clin Epigenetics, 2016. **8**: p. 37.
18. Condliffe, D., et al., *Cross-region reduction in 5-hydroxymethylcytosine in Alzheimer's disease brain.* Neurobiol Aging, 2014. **35**(8): p. 1850-4.
19. Deissler, H., et al., *Rapid protein sequencing by tandem mass spectrometry and cDNA cloning of p20-CGGBP. A novel protein that binds to the unstable triplet repeat 5'-d(CGG)n-3' in the human FMR1 gene.* J Biol Chem, 1997. **272**(27): p. 16761-8.
20. Tsai, R.Y., *p53-guided response to nucleostemin loss in normal versus cancer cells.* Cell Death Dis, 2015. **6**: p. e2030.

21. Caricasole, A., et al., *Induction of Dickkopf-1, a negative modulator of the Wnt pathway, is associated with neuronal degeneration in Alzheimer's brain.* J Neurosci, 2004. **24**(26): p. 6021-7.
22. Toledo, E.M. and N.C. Inestrosa, *Activation of Wnt signaling by lithium and rosiglitazone reduced spatial memory impairment and neurodegeneration in brains of an APPswe/PSEN1DeltaE9 mouse model of Alzheimer's disease.* Mol Psychiatry, 2010. **15**(3): p. 272-85, 228.
23. Khare, T., et al., *5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary.* Nat Struct Mol Biol, 2012. **19**(10): p. 1037-43.
24. Slieker, R.C., et al., *Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array.* Epigenetics Chromatin, 2013. **6**(1): p. 26.
25. Pedersen, B.S., et al., *Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values.* Bioinformatics, 2012. **28**(22): p. 2986-8.
26. Esiri, M.M., *Brain banks: the Oxford experience.* J Neural Transm Suppl, 1993. **39**: p. 25-30.
27. Mastroeni, D., et al., *ANK1 is up-regulated in laser captured microglia in Alzheimer's brain; the importance of addressing cellular heterogeneity.* PLOS ONE, 2017. **12**(7): p. e0177814.
28. Voronin, D.A. and E.V. Kiseleva, *Functional role of proteins containing ankyrin repeats.* Cell and Tissue Biology, 2008. **2**(1): p. 1-12.
29. Zhao, J., et al., *A genome-wide profiling of brain DNA hydroxymethylation in Alzheimer's disease.* Alzheimers Dement, 2017. **13**(6): p. 674-688.
30. Ellison, E.M., M.A. Bradley-Whitman, and M.A. Lovell, *Single-Base Resolution Mapping of 5-Hydroxymethylcytosine Modifications in Hippocampus of Alzheimer's Disease Subjects.* J Mol Neurosci, 2017.
31. Smith, R.G., et al., *Elevated DNA methylation across a 48kb region spanning the HOXA gene cluster on chromosome 7 is associated with Alzheimer's disease neuropathology in the prefrontal cortex and superior temporal gyrus.* Alzheimers Dement, 2017. **[Under Review]**.
32. Smith, A.R., et al., *Increased DNA methylation near TREM2 is consistently seen in the superior temporal gyrus in Alzheimer's disease brain.* Neurobiol Aging, 2016. **47**: p. 35-40.
33. R Development Core Team, *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria 2012, 2012.
34. Gentleman, R.C., et al., *Bioconductor: open software development for computational biology and bioinformatics.* Genome Biol, 2004. **5**(10): p. R80.
35. Davis, S., et al., *Methylumi: Handle Illumina Methylation Data 2012..*
. R package version 2.2.0, 2012.
36. Chen, Y.A., et al., *Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray.* Epigenetics, 2013. **8**(2): p. 203-9.
37. Pidsley, R., et al., *A data-driven approach to preprocessing Illumina 450K methylation array data.* BMC Genomics, 2013. **14**: p. 293.
38. Guintivano, J., M. Aryee, and Z. Kaminsky, *A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression.* Epigenetics, 2013. **8**(3).
39. Phipson, B., J. Maksimovic, and A. Oshlack, *missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform.* Bioinformatics, 2016. **32**(2): p. 286-8.

TABLES

Table 1: DMoPs, DMPs and DHPs with a neuropathology-associated difference in levels above our significance threshold ($P < 5.0 \times 10^{-5}$) in more than one comparison. Shown for each probe are chromosomal location (hg19), Gene annotation (UCSC/GREAT), with corrected effect size (difference (Δ) between Braak score 0 and Braak score VI) and corresponding *P* value after adjusting for the covariates of age and sex, and an additional *P* value after correcting for neuron/glia proportions (*P* CETS corrected) for total DNA modifications (BS), 5mC (OxBS) and 5hmC (BS – OxBS), with annotation as to whether the mean 5hmC level was >0. All P values <0.05 are shown in bold.

			DMoP (BS)			DMP (OxBS)			DHP (BS - OxBS)			
Gene name	**Probe**	**Position**	***P***	**Δ**	***P* CETS corrected**	***P***	**Δ**	***P* CETS corrected**	***P***	**Δ**	***P* CETS corrected**	***5hmC* Positive probe**
KCNJ9	cg20066612	1:160050948	**6.79×10^{-06}**	**3.25**	**2.73×10^{-05}**	**0.027**	**-2.89**	**0.033**	**2.70×10^{-05}**	**5.84**	**6.47×10^{-05}**	✓
H3F3A	cg15624314	1:226251003	0.32	-0.25	0.45	**6.14×10^{-07}**	**-1.84**	**1.68×10^{-06}**	**4.30×10^{-05}**	**1.59**	**9.50×10^{-05}**	
OPN3	cg08240832	1:241791729	0.129	-1.81	0.097	**3.21×10^{-05}**	**7.5**	**8.12×10^{-05}**	**7.35×10^{-06}**	**-10.09**	**1.99×10^{-05}**	✓
TMEM53	cg13876163	1:45140495	0.145	0.33	0.178	**2.73×10^{-05}**	**-2.25**	**4.28×10^{-05}**	**8.39×10^{-06}**	**2.55**	**1.45×10^{-05}**	
NHEJ1	cg18070033	2:220025364	0.919	0.02	0.848	**3.64×10^{-06}**	**-1.84**	**8.65×10^{-06}**	**6.24×10^{-06}**	**1.84**	**1.38×10^{-05}**	
ACVR2B	cg17026303	3:38495161	0.948	-0.02	0.976	**6.70×10^{-07}**	**-2.46**	**1.84×10^{-06}**	**2.74×10^{-06}**	**2.39**	**7.50×10^{-06}**	
ROBO2	cg16933762	3:77147109	0.493	-1.12	0.386	**2.77×10^{-05}**	**10.36**	**7.41×10^{-05}**	**2.80×10^{-05}**	**-10.61**	**7.11×10^{-05}**	
CGGBP1	cg15147060	3:88108213	0.44	-0.13	0.536	**3.65×10^{-09}**	**-1.92**	**1.01×10^{-08}**	**3.94×10^{-07}**	**1.73**	**1.07×10^{-06}**	
RAPGEF6	cg05052335	5:130970657	0.215	-0.17	0.217	**9.73×10^{-07}**	**-1.56**	**2.38×10^{-06}**	**6.00×10^{-06}**	**1.39**	**1.48×10^{-05}**	
GFRA1	cg03503087	10:118032626	0.731	0.06	0.765	**2.38×10^{-06}**	**-1.89**	**4.50×10^{-06}**	**1.02×10^{-05}**	**1.97**	**1.79×10^{-05}**	
FHOD1	cg11530914	16:67281528	0.642	0.12	0.68	**2.86×10^{-05}**	**-1.87**	**7.39×10^{-05}**	**7.74×10^{-06}**	**1.95**	**2.09×10^{-05}**	
FOXF1	cg05341384	16:86331661	0.875	-0.1	0.948	**4.15×10^{-06}**	**-4.92**	**1.10×10^{-05}**	**2.01×10^{-05}**	**4.41**	**4.84×10^{-05}**	
KCNC3	cg07919466	19:50833956	0.73	-0.09	0.986	**8.41×10^{-07}**	**-1.83**	**2.28×10^{-06}**	**1.03×10^{-05}**	**1.69**	**1.68×10^{-05}**	
APP	cg12144476	21:27541894	0.659	0.08	0.562	**4.85×10^{-06}**	**-1.39**	**9.79×10^{-06}**	**1.61×10^{-05}**	**1.44**	**4.12×10^{-05}**	

Table 2: Identification of multiple-probe DMRs associated with Braak stage. Spatially correlated *P* values in a 500bp sliding window were identified using *comb-p* [25]. Shown are those with a Sidak-corrected *P* value > 0.05.

Position	Gene Name	Probes in DMR	*Comb-p P* Value	*Comb-p* Sidak corrected *P* Value
chr3:55517806-55518442	*WNT5A*	8	2.53×10^{-9}	3.17×10^{-4}
chr14:103367306-103367592	*TRAF3*	3	8.12×10^{-8}	0.022
chr8:41519308-41519400	*ANK1*	2	2.89×10^{-8}	0.025
chr10:63809073-63809171	*ARID5B*	6	3.60×10^{-8}	0.029

Figure 1

A

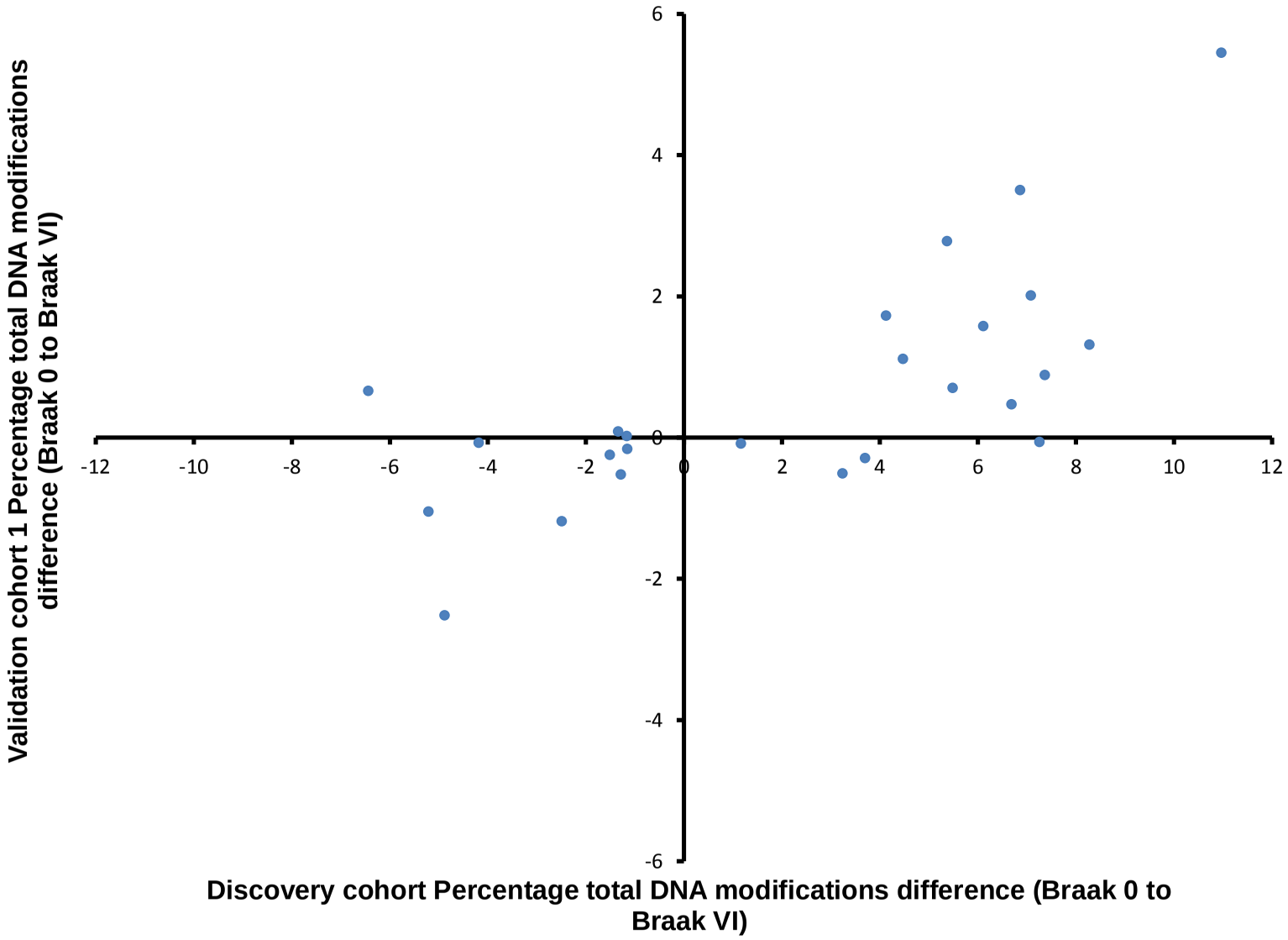

Validation cohort 1 Percentage total DNA modifications difference (Braak 0 to Braak VI)
Discovery cohort Percentage total DNA modifications difference (Braak 0 to Braak VI)
6
4
2
0
-2
-4
-6
-12
-10
-8
-6
-4
-2
0
2
4
6
8
10
12

Figure 1

B

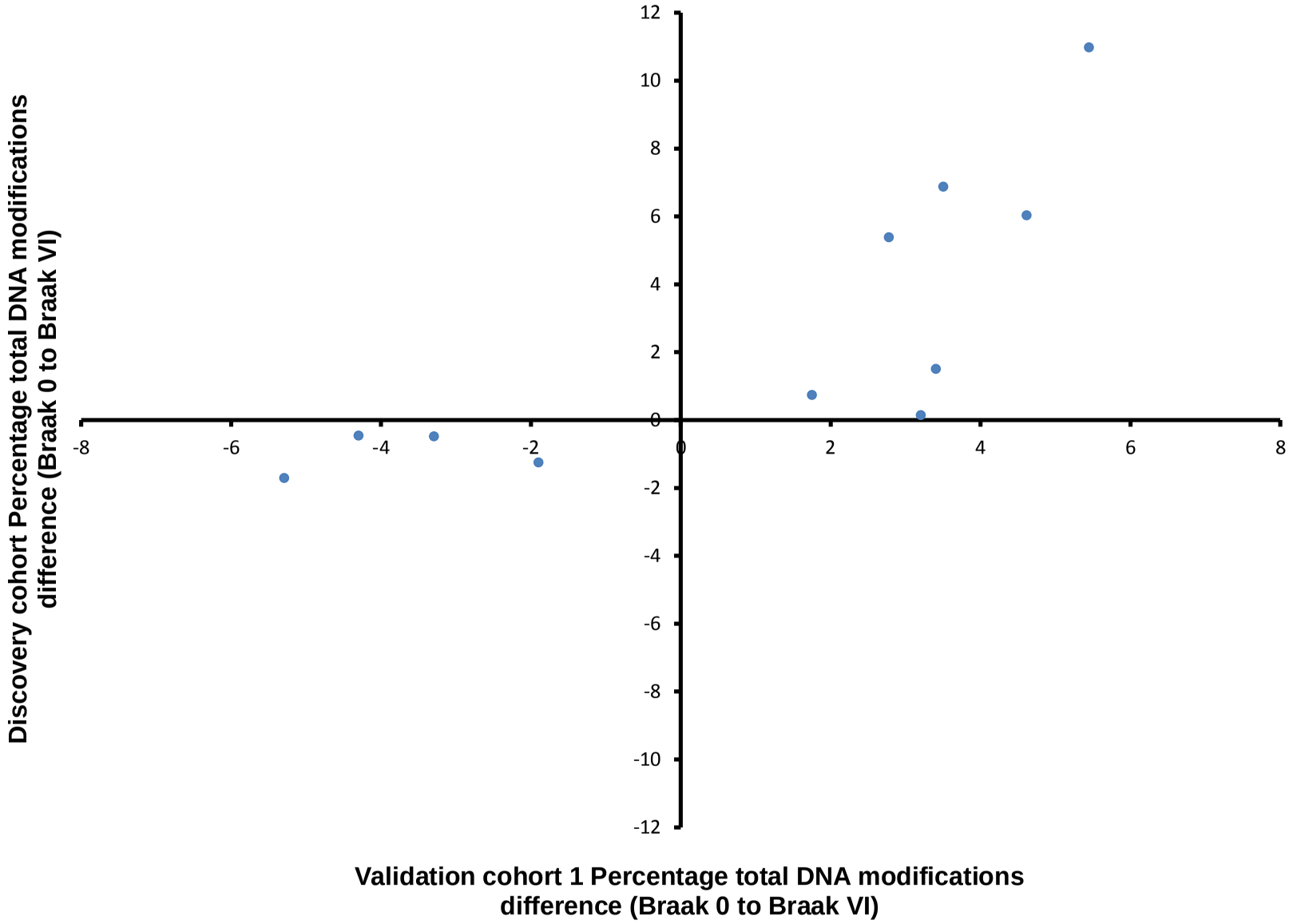

Discovery cohort Percentage total DNA modifications difference (Braak 0 to Braak VI)
Validation cohort 1 Percentage total DNA modifications difference (Braak 0 to Braak VI)
12
10
8
6
4
2
0
-2
-4
-6
-8
-10
-12
-8
-6
-4
-2
0
2
4
6
8

Figure 2

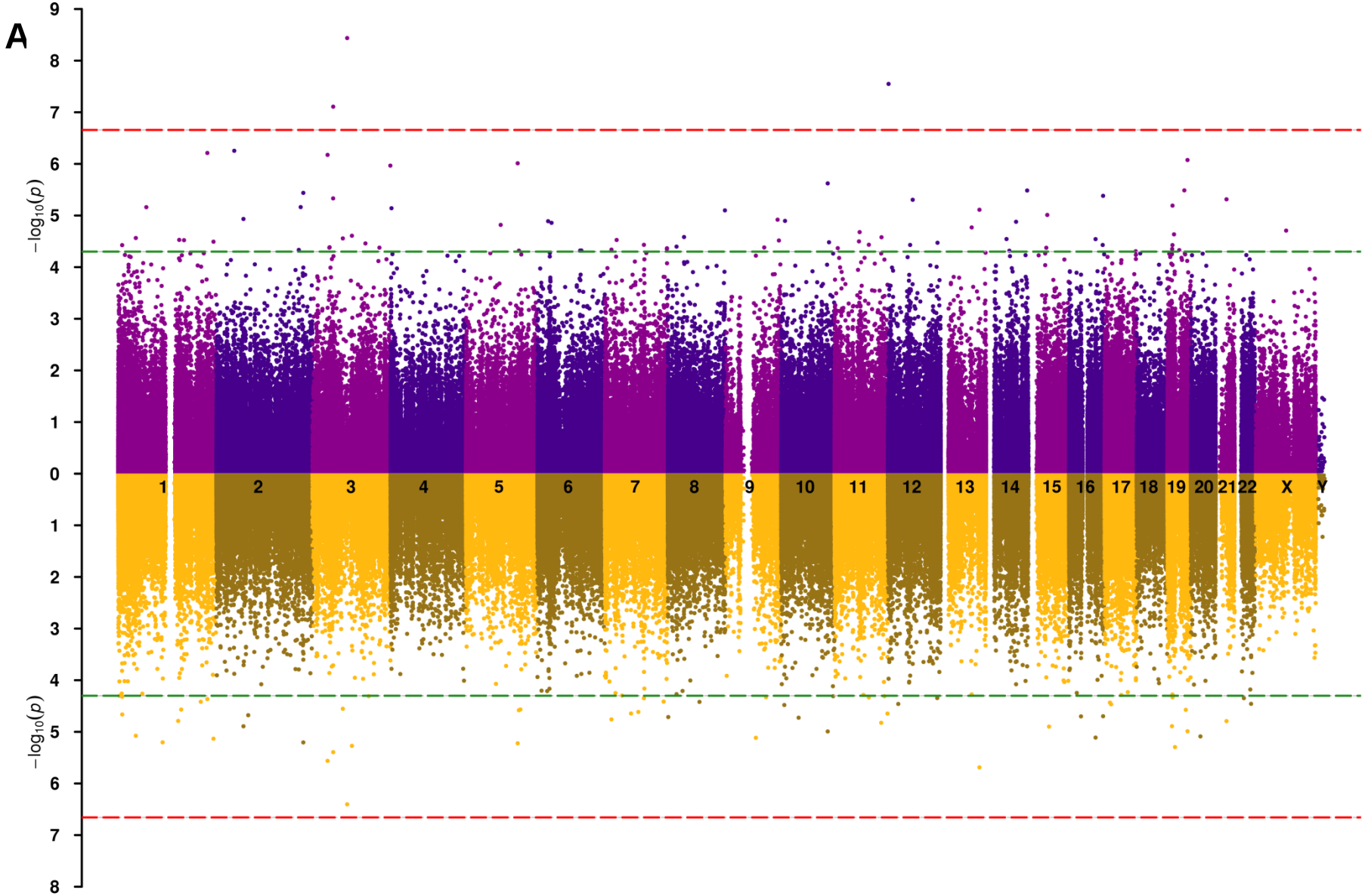

A
9
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
$-\log_{10}(p)$
$-\log_{10}(p)$

Figure 2

B

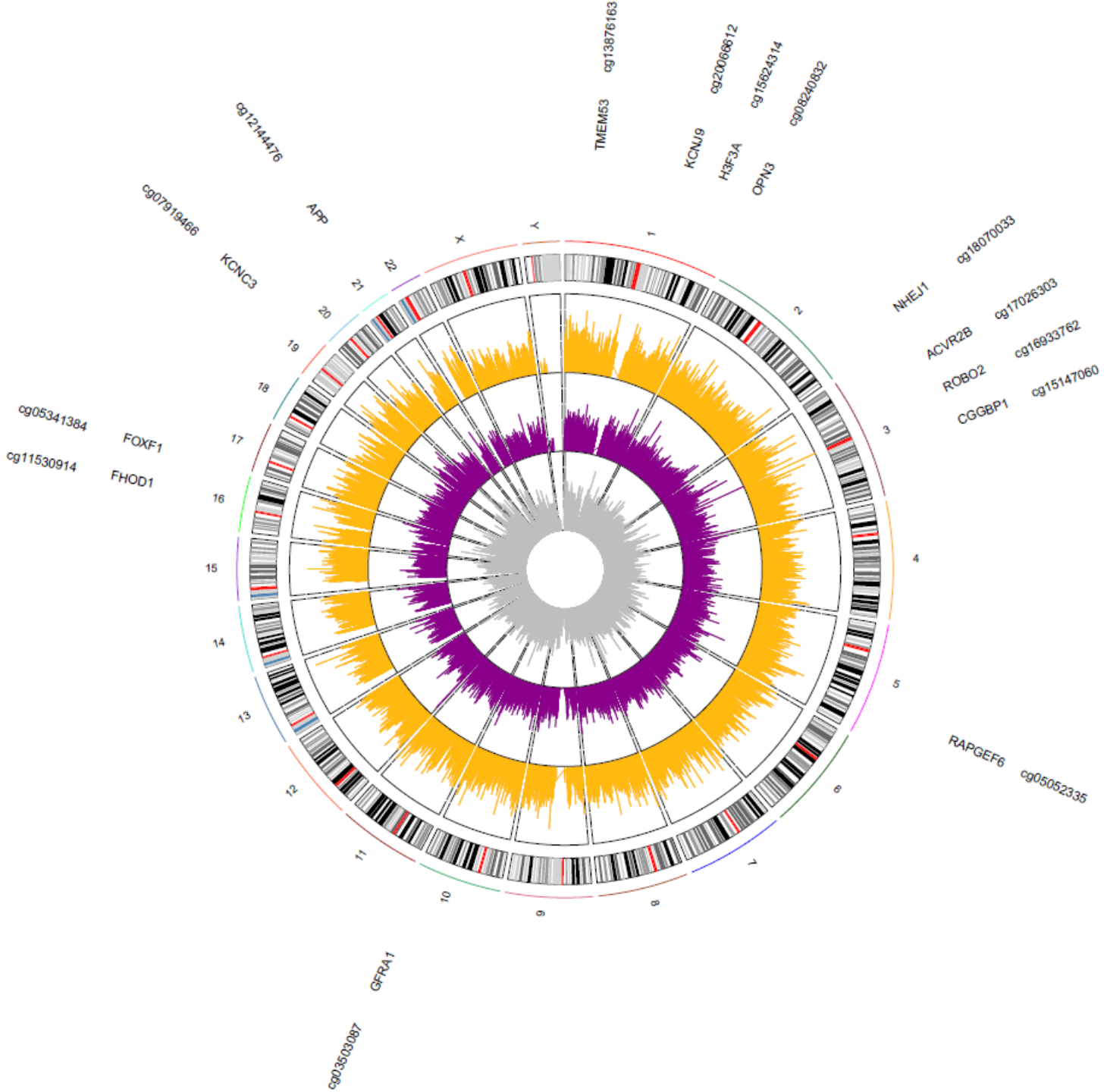

cg13876163
TMEM53
cg20066612
KCNJ9
cg15624314
H3F3A
cg08240832
OPN3
cg12144476
APP
cg07919466
KCNC3
cg18070033
NHEJ1
cg17026303
ACVR2B
cg16933762
ROBO2
cg15147060
CGGBP1
cg05341384
FOXF1
cg11530914
FHOD1
RAPGEF6
cg05052335
GFRA1
cg03503087
X
Y
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

Figure 2

C

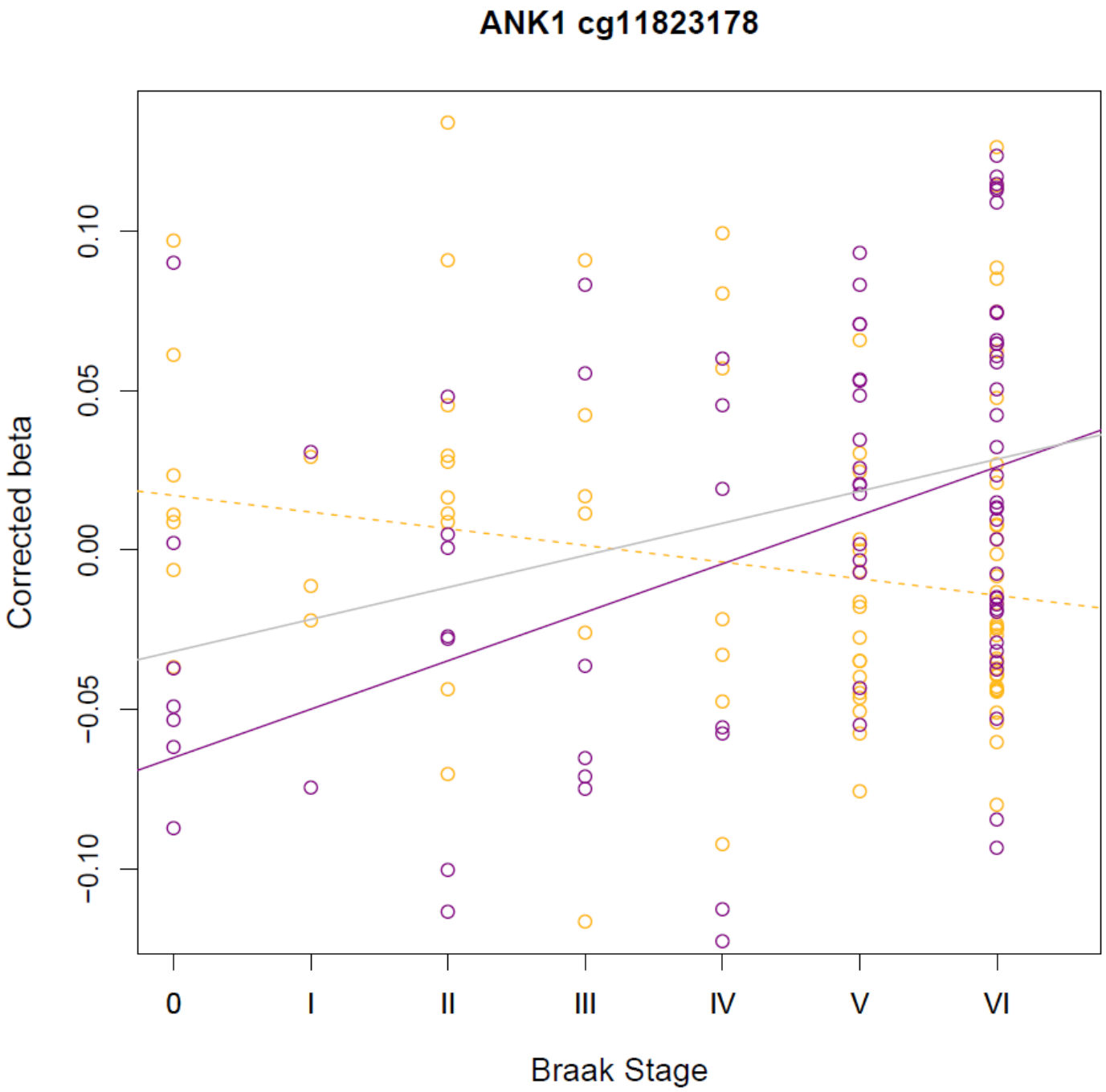

ANK1 cg11823178
Corrected beta
0.10
0.05
0.00
−0.05
−0.10
0
I
II
III
IV
V
VI
Braak Stage

Figure 2

D

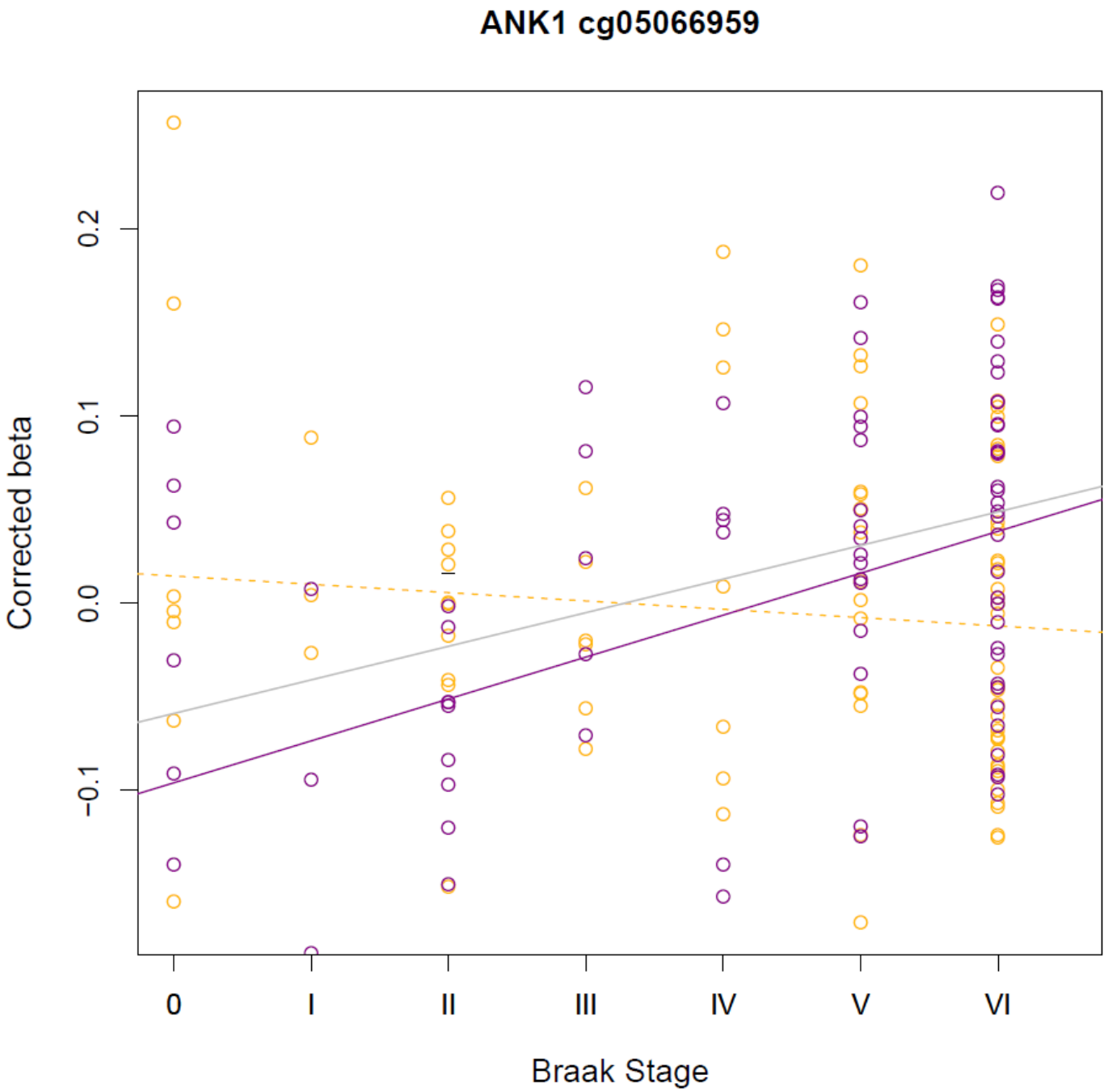

ANK1 cg05066959
Corrected beta
0.2
0.1
0.0
−0.1
0
I
II
III
IV
V
VI
Braak Stage

Figure 3

A

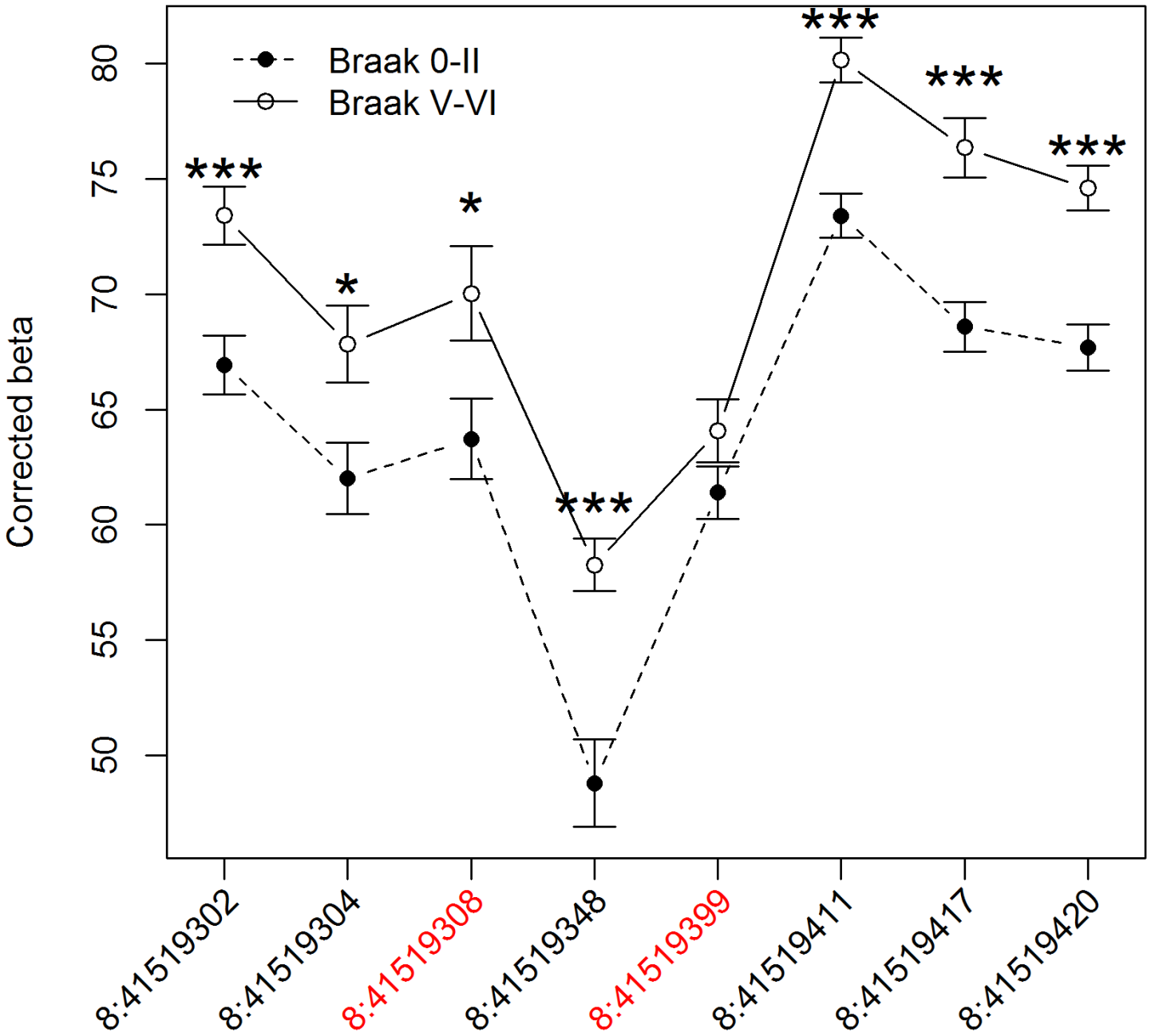

Braak 0-II
Braak V-VI
Corrected beta
50
55
60
65
70
75
80
***
*
*
***
***
***
8:41519302
8:41519304
8:41519308
8:41519348
8:41519399
8:41519411
8:41519417
8:41519420

Figure 3

B

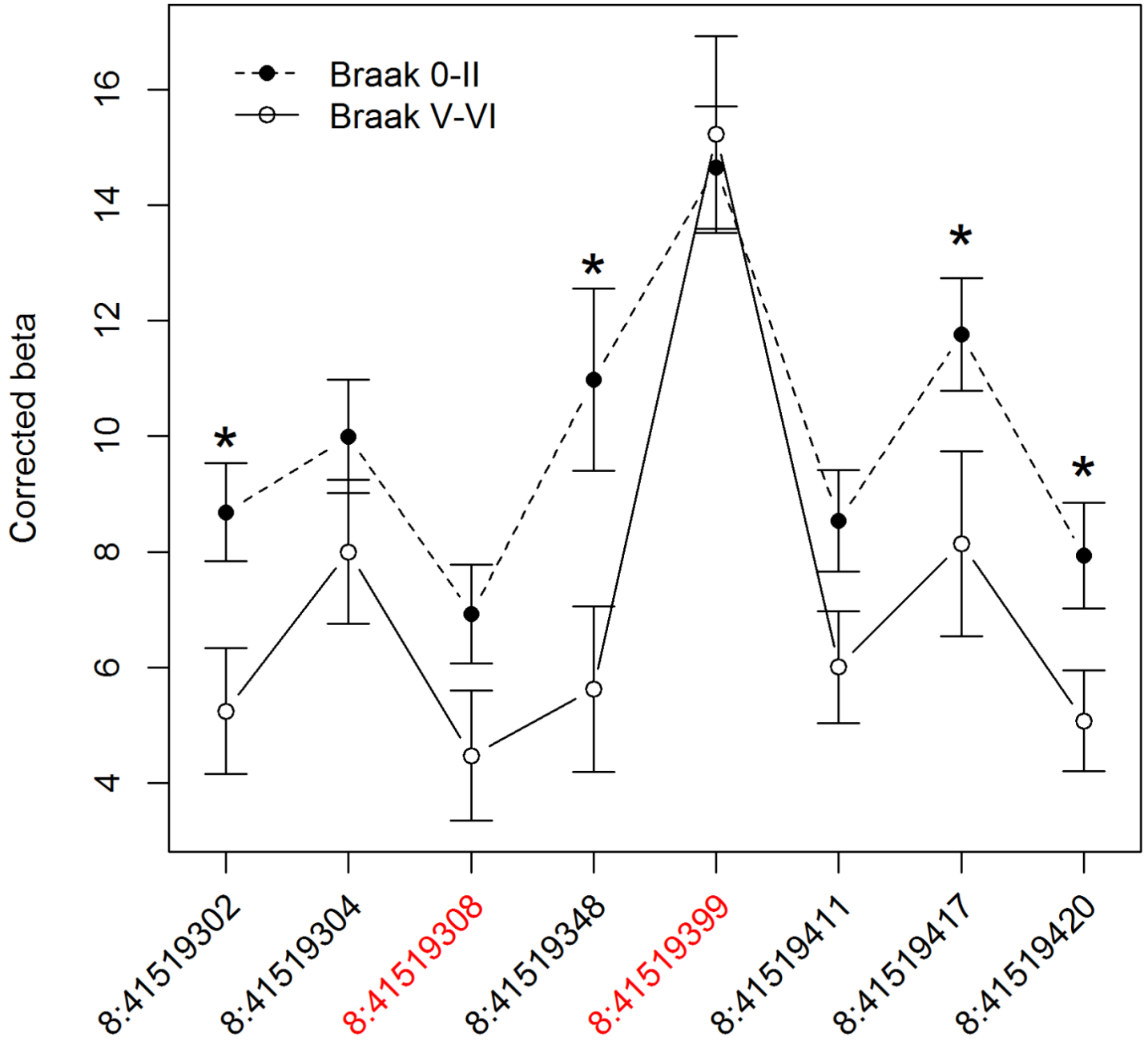

Braak 0-II
Braak V-VI
Corrected beta
4
6
8
10
12
14
16
*
*
*
*
8:41519302
8:41519304
8:41519308
8:41519348
8:41519399
8:41519411
8:41519417
8:41519420

Figure 3

C

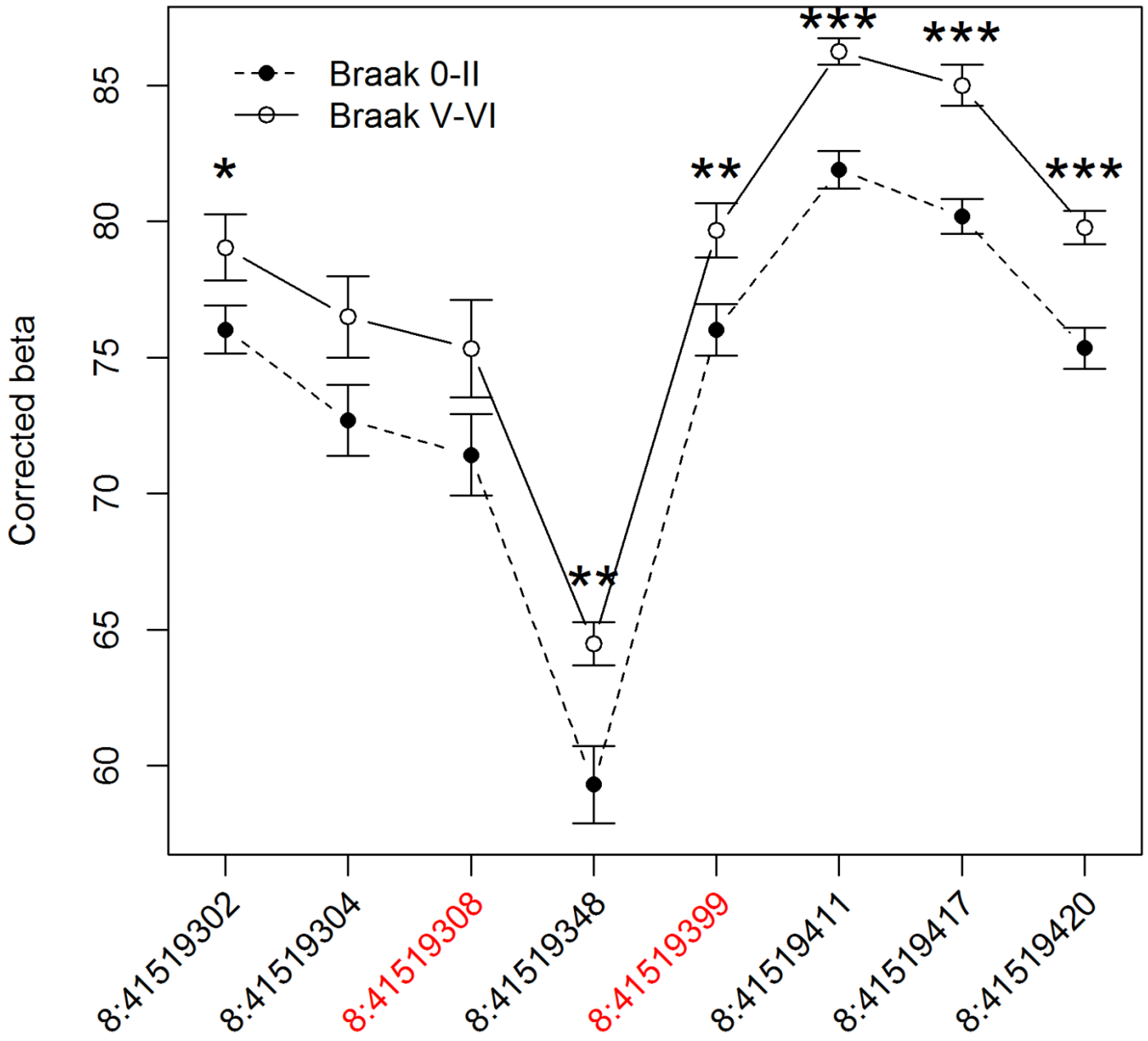
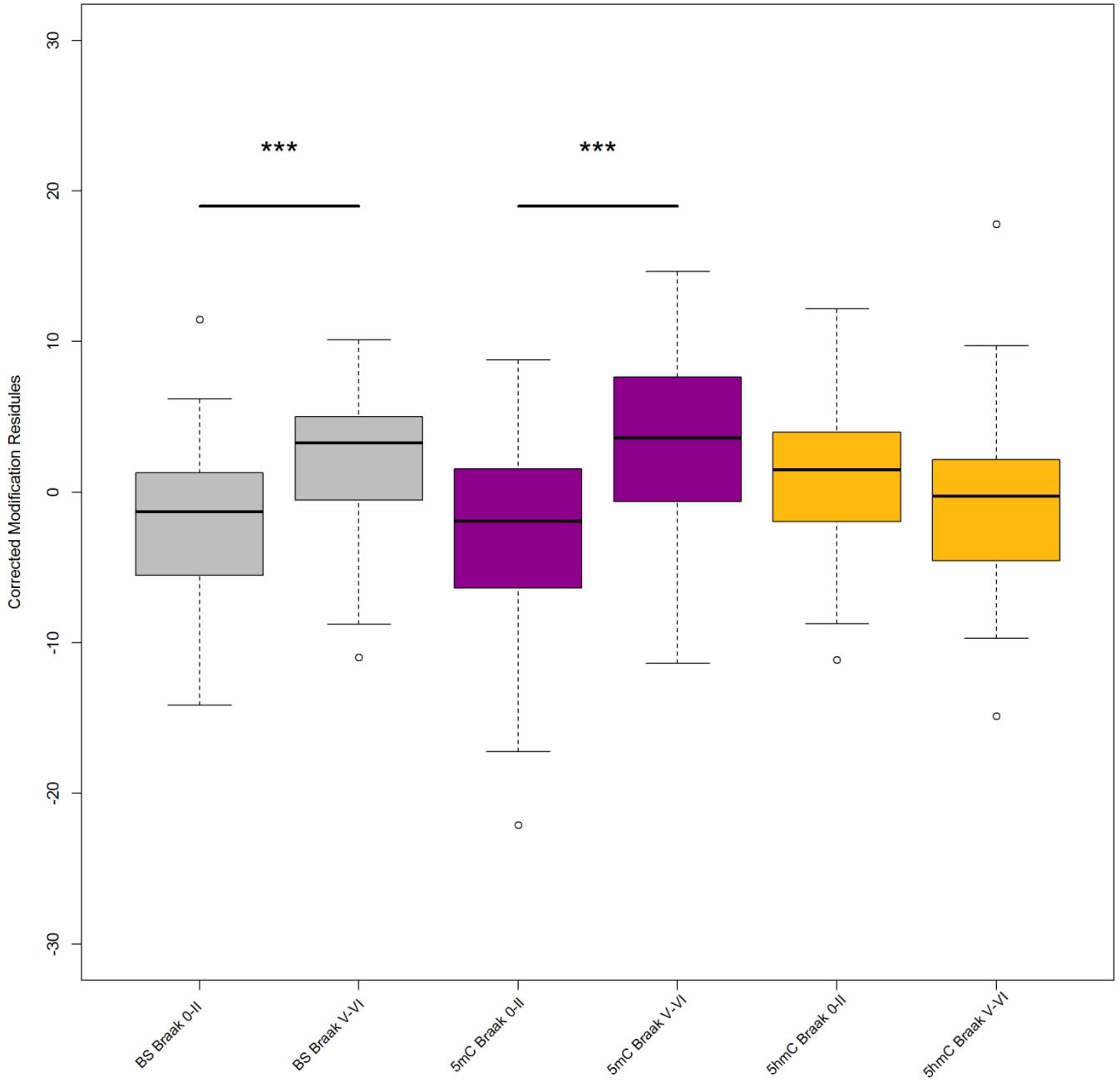

Braak 0-II
Braak V-VI
Corrected beta
60
65
70
75
80
85
*
**
**
***
***
***
8:41519302
8:41519304
8:41519308
8:41519348
8:41519399
8:41519411
8:41519417
8:41519420

Figure 3

D


Corrected Modification Residules
30
20
10
0
-10
-20
-30
***
***
BS Braak 0-II
BS Braak V-VI
5mC Braak 0-II
5mC Braak V-VI
5hmC Braak 0-II
5hmC Braak V-VI

Supplementary Table 1: Significant neuropathology-associated differentially modified positions (DMoPs) in the EC. Shown for each probe are chromosomal location (hg19), UCSC annotation, GREAT annotation, with corrected effect size (difference (Δ) in total DNA modification levels) and corresponding P value between Braak score 0 and Braak score VI after adjusting for the covariates of age and sex, and an additional corresponding P value also correcting for neuron/glia proportions (P CETS corrected). Shown for all probes are Δ and corresponding P value for 5mC and 5hmC levels alone, with annotation as to whether the mean 5hmC level was >0. All P values <0.05 are shown in bold.

Probe information					Total modifications (BS)			5mC (OxBS)		5hmC (BS - OxBS)		
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	*P*	Δ	*P*	Δ	5hmC Positive probe
1	cg21774827	19:54379120	MYADM	PRKCG	**3.46E-07**	7.37	**4.76E-07**	0.066	4.52	0.264	2.68	✓
2	cg26925343	16:28857904	TUFM	SH2B1	**4.06E-06**	-1.28	**2.98E-06**	**0.042**	-1.27	0.998	0.00	
3	cg17554875	5:179520998	-	RNF130	**4.24E-06**	7.26	**1.77E-05**	0.598	0.83	**8.75E-04**	6.23	✓
4	cg27642528	10:88814651	GLUD1	GLUD1	**5.75E-06**	3.71	**1.89E-05**	0.825	0.28	**0.005**	3.55	✓
5	cg20066612	1:160050948	KCNJ9	KCNJ9	**6.79E-06**	3.25	**2.73E-05**	**0.027**	-2.89	**2.70E-05**	5.84	✓
6	cg03975188	7:148131599	-	CUL1	**7.00E-06**	8.28	**2.07E-05**	0.867	0.33	**0.002**	8.39	✓
7	cg17304222	5:180076905	FLT4	FLT4	**8.29E-06**	-1.16	**1.71E-05**	**0.002**	-1.07	0.860	-0.07	
8	cg05066959	8:41519308	ANK1	NKX6-3	**8.37E-06**	10.98	**2.94E-05**	**8.03E-05**	13.71	0.349	-2.72	✓
9	cg01042637	3:130743723	ASTE1	ASTE1	**9.75E-06**	6.12	**3.20E-05**	0.172	2.65	0.112	3.39	✓
10	cg04658038	17:64800166	PRKCA	CACNG5	**1.08E-05**	4.48	**1.32E-05**	0.634	0.87	0.068	3.72	✓
11	cg06653632	12:129281444	SLC15A4	SLC15A4	**1.21E-05**	6.87	**4.47E-05**	0.151	3.00	0.083	3.81	✓
12	cg04904331	7:49813033	VWC2	VWC2	**1.59E-05**	-1.15	**2.51E-05**	0.170	0.43	**2.89E-04**	-1.54	
13	cg21151057	7:157393764	PTPRN2	DNAJB6	**1.59E-05**	5.49	**5.19E-05**	0.983	0.04	**0.027**	5.54	✓
14	cg17867333	5:178423163	GRM6	GRM6	**1.87E-05**	-5.21	**3.08E-05**	0.148	-1.94	**0.007**	-3.30	
15	cg13851211	16:50321678	ADCY7	ADCY7	**2.91E-05**	7.09	**2.87E-05**	**0.034**	4.13	0.224	2.86	✓
16	cg15197657	5:176916434	PDLIM7	PDLIM7	**3.28E-05**	-6.43	**9.83E-05**	0.183	-5.18	0.820	-0.84	
17	cg20618448	19:49962324	ALDH16A1	ALDH16A1	**3.29E-05**	5.38	**3.33E-05**	**0.001**	5.94	0.721	-0.44	✓
18	cg06898199	7:128502890	ATP6V1F	ATP6V1F	**3.44E-05**	1.17	**1.69E-05**	**0.017**	0.68	0.154	0.47	
19	cg23022785	20:10153355	-	SNAP25	**3.68E-05**	-1.50	**6.06E-05**	0.375	-0.52	0.124	-1.02	
20	cg06192381	3:50305066	SEMA3B	LSMEM2	**4.26E-05**	-1.33	**1.54E-04**	0.316	0.43	**0.001**	-1.62	
21	cg00627837	22:19133638	DGCR14	DGCR14	**4.32E-05**	-4.18	**1.14E-04**	0.431	-0.94	**0.033**	-3.11	
22	cg26040816	20:42816135	JPH2	JPH2	**4.62E-05**	-2.49	**1.03E-04**	0.787	0.28	**0.010**	-2.59	
23	cg01348086	5:176785296	RGS14	RGS14	**4.73E-05**	4.13	**8.67E-05**	**4.87E-04**	3.08	0.163	1.10	✓
24	cg03309308	16:68525329	-	SMPD3	**4.78E-05**	-4.88	**1.13E-04**	**0.009**	-3.08	0.277	-1.61	
25	cg10794439	8:41518051	ANK1;MIR486	NKX6-3	**4.81E-05**	6.69	**1.98E-04**	**0.022**	6.70	0.901	0.32	✓

Supplementary Table 2: Significant neuropathology-associated differentially modified positions (DMoPs) in the CER. Shown for each probe are chromosomal location (hg19), UCSC annotation, GREAT annotation, with corrected effect size (difference (Δ) in total DNA modification levels) and corresponding P value between Braak score 0 and Braak score VI after adjusting for the covariates of age and sex, and an additional corresponding P value also correcting for neuron/glia proportions (P CETS corrected). Shown for all probes are Δ and corresponding P value for 5mC and 5hmC levels alone, with annotation as to whether the mean 5hmC level was >0. All P values <0.05 are shown in bold.

Probe information					**Total DNA modifications (BS)**			**5mC (OxBS)**		**5hmC (BS - OxBS)**		
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	*P*	Δ	*P*	Δ	5hmC Positive probe
1	cg26572042	1:21113758	HP1BP3	HP1BP3	**1.45E-06**	-1.00	**6.13E-07**	**0.001**	-0.62	0.153	-0.38	
2	cg17598713	6:33265534	RGL2	RGL2	**4.83E-06**	-2.25	**2.08E-06**	0.279	-0.78	0.073	-1.48	
3	cg18107396	6:35436198	RPL10A	RPL10A	**1.02E-05**	0.56	**1.37E-05**	0.142	0.20	**0.050**	0.36	
4	cg16085178	6:135818824	AHI1;C6orf217	AHI1	**1.49E-05**	0.67	**2.03E-05**	0.798	0.07	**0.049**	0.61	
5	cg03415001	4:181480012	-	-	**2.55E-05**	-3.42	**2.27E-05**	0.537	0.53	**0.001**	-3.89	✓
6	cg04074793	10:129924374	MKI67	MKI67	**2.57E-05**	-0.72	**3.33E-05**	0.841	-0.04	**0.009**	-0.68	
7	cg16720675	18:19284817	ABHD3	ABHD3	**3.07E-05**	0.96	**3.50E-05**	0.927	-0.05	**0.041**	1.00	
8	cg02215070	7:134144055	AKR1B1	AKR1B1	**3.59E-05**	0.79	**3.66E-05**	0.407	0.15	**0.009**	0.64	✓
9	cg04899175	12:117798748	NOS1	NOS1	**3.87E-05**	0.92	**5.12E-05**	**0.006**	1.07	0.708	-0.15	
10	cg08161720	3:52188073	WDR51A	POC1A	**3.89E-05**	-1.41	**4.18E-05**	0.377	-0.44	0.101	-0.95	
11	cg14289985	19:57018848	ZNF471	ZNF471	**4.17E-05**	1.07	**5.65E-05**	0.161	-0.44	**2.61E-05**	1.53	✓
12	cg24022821	6:28831886	-	TRIM27	**4.35E-05**	0.89	**5.40E-05**	**0.036**	0.50	0.144	0.40	✓
13	cg07612045	7:144533234	TPK1	TPK1	**4.89E-05**	-0.79	**5.70E-05**	0.899	-0.05	0.126	-0.74	

Supplementary Table 3: Significant neuropathology-associated differentially methylated positions (DMPs) in the EC. Shown for each probe are chromosomal location (hg19), UCSC annotation, GREAT annotation, with corrected effect size (difference (Δ) in 5mC levels) and corresponding P value between Braak score 0 and Braak score VI after adjusting for the covariates of age and sex, and an additional corresponding P value also correcting for neuron/glia proportions (P CETS corrected). Shown for all probes are Δ and corresponding P value for total DNA modifications and 5hmC levels alone, with annotation as to whether the mean 5hmC level was >0. All P values <0.05 are shown in bold.

Probe information					5mC (OxBS)			Total modifications (BS)		5hmC (BS - OxBS)		
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	*P*	Δ	*P*	Δ	5hmC Positive probe
1	cg15147060	3:88108213	CGGBP1	CGGBP1	**3.65E-09**	-1.92	**1.01E-08**	0.440	-0.13	**3.94E-07**	1.73	
2	cg10696062	12:1726028	WNT5B	WNT5B	**2.82E-08**	13.73	**7.60E-08**	**0.009**	4.45	**1.07E-04**	-9.58	
3	cg00106685	3:52720133	GNL3;PBRM1	GNL3	**7.80E-08**	-1.45	**1.89E-07**	0.413	-0.13	**4.04E-06**	1.39	
4	cg18110359	2:45179970	-	SIX3	**5.57E-07**	-2.18	**1.51E-06**	0.150	-0.60	**1.33E-04**	1.95	
5	cg15624314	1:226251003	H3F3A;LOC440926	H3F3A	**6.14E-07**	-1.84	**1.68E-06**	0.320	-0.25	**4.30E-05**	1.59	
6	cg17026303	3:38495161	LOC100128640;ACVR2B	ACVR2B	**6.70E-07**	-2.46	**1.84E-06**	0.948	-0.02	**2.74E-06**	2.39	
7	cg07919466	19:50833956	KCNC3	KCNC3	**8.41E-07**	-1.83	**2.28E-06**	0.730	-0.09	**1.03E-05**	1.69	
8	cg05052335	5:130970657	RAPGEF6	RAPGEF6	**9.73E-07**	-1.56	**2.38E-06**	0.215	-0.17	**6.00E-06**	1.39	
9	cg26332552	3:197402549	MIR922;KIAA0226	KIAA0226	**1.08E-06**	11.17	**2.95E-06**	**4.17E-04**	6.57	**0.037**	-4.40	✓
10	cg03503087	10:118032626	GFRA1	GFRA1	**2.38E-06**	-1.89	**4.50E-06**	0.731	0.06	**1.02E-05**	1.97	
11	cg02456292	19:42817262	TMEM145	TMEM145	**3.26E-06**	-2.72	**6.15E-06**	0.147	-0.49	**7.63E-04**	2.22	
12	cg04418576	14:103851554	MARK3	MARK3	**3.28E-06**	-1.67	**8.53E-06**	0.218	-0.23	**1.86E-04**	1.41	
13	cg18070033	2:220025364	NHEJ1	NHEJ1	**3.64E-06**	-1.84	**8.65E-06**	0.919	0.02	**6.24E-06**	1.84	
14	cg05341384	16:86331661	-	FOXF1	**4.15E-06**	-4.92	**1.10E-05**	0.875	-0.10	**2.01E-05**	4.41	
15	cg14467668	3:52571040	LOC440957	SMIM4	**4.66E-06**	-1.38	**1.22E-05**	0.874	0.03	**1.94E-04**	1.34	
16	cg12144476	21:27541894	APP	APP	**4.85E-06**	-1.39	**9.79E-06**	0.659	0.08	**1.61E-05**	1.44	
17	cg12366118	12:62860687	MON2	MON2	**4.96E-06**	-1.73	**1.35E-05**	0.126	-0.31	**2.07E-04**	1.39	
18	cg14294321	19:12895288	-	JUNB	**6.43E-06**	7.16	**1.57E-05**	0.274	1.45	**0.005**	-5.58	✓
19	cg00639289	2:213403734	ERBB4	ERBB4	**6.86E-06**	-1.32	**1.62E-05**	0.418	-0.11	**5.12E-04**	1.11	
20	cg24516830	1:71694460	-	ZRANB2	**6.89E-06**	9.60	**1.66E-05**	0.122	2.57	**0.003**	-6.74	✓
21	cg00847044	4:2419822	ZFYVE28	ZFYVE28	**7.22E-06**	-1.82	**1.93E-05**	0.974	-0.01	**1.66E-04**	1.77	
22	cg25329939	13:98085947	RAP2A	RAP2A	**7.73E-06**	-1.29	**1.74E-05**	0.074	0.31	**2.05E-06**	1.56	
23	cg03792042	8:145045235	PLEC1	PARP10	**7.95E-06**	10.13	**2.10E-05**	0.162	3.03	**0.002**	-6.85	✓
24	cg06624337	15:47477411	-	SEMA6D	**9.74E-06**	-1.64	**2.63E-05**	0.460	-0.17	**0.001**	1.41	
25	cg03263685	2:68480160	PPP3R1	PPP3R1	**1.17E-05**	-1.23	**3.08E-05**	0.932	0.01	**1.28E-05**	1.21	
26	cg01888395	9:132145105	-	NTMT1	**1.20E-05**	7.14	**2.38E-05**	0.068	3.10	**0.017**	-4.13	✓
27	cg22782712	10:9920178	-	-	**1.27E-05**	5.39	**2.27E-05**	0.477	-0.93	**1.82E-04**	-6.71	
28	cg05645927	6:27344393	ZNF204P	ZNF391	**1.29E-05**	7.21	**6.68E-06**	0.793	0.40	**0.001**	-6.64	
29	cg14102251	14:75745098	FOS	FOS	**1.32E-05**	-1.67	**3.51E-05**	0.487	-0.15	**8.36E-05**	1.50	
30	cg02412050	6:35995429	MAPK1	MAPK14	**1.39E-05**	-1.41	**3.63E-05**	**0.031**	-0.39	**0.007**	0.94	
31	cg24631913	5:87971905	LOC645323	MEF2C	**1.52E-05**	-1.41	**3.62E-05**	0.598	-0.12	**5.80E-04**	1.25	
32	cg23648516	13:78467552	-	EDNRB	**1.70E-05**	9.69	**4.27E-05**	0.673	-0.63	**5.31E-05**	-10.48	✓
33	cg21902245	X:79270180	TBX22	TBX22	**1.96E-05**	-6.42	**4.51E-05**	**0.017**	-4.85	0.420	1.29	✓

Probe information					5mC (OxBS)			Total modifications (BS)		5hmC (BS - OxBS)		
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	*P*	Δ	*P*	Δ	5hmC Positive probe
34	cg16898576	11:63449499	RTN3	RTN3	**2.09E-05**	-1.31	**5.45E-05**	0.512	-0.11	**7.91E-04**	1.15	
35	cg15235057	19:16683989	SLC35E1	SLC35E1	**2.31E-05**	-1.46	**3.52E-04**	0.522	-0.12	**2.96E-04**	1.26	
36	cg03160445	3:100120349	LNP1;TOMM70A	TOMM70A	**2.46E-05**	-1.48	**6.34E-05**	0.680	0.07	**5.35E-06**	1.50	
37	cg11823178	8:41519399	ANK1;MIR486	NKX6-3	**2.61E-05**	9.31	**6.64E-05**	**4.68E-04**	6.03	0.068	-3.20	✓
38	cg11507178	11:118781515	BCL9L	BCL9L	**2.63E-05**	-1.92	**7.07E-05**	0.219	-0.59	**0.039**	1.22	
39	cg13876163	1:45140495	C1orf228;TMEM53	TMEM53	**2.73E-05**	-2.25	**4.28E-05**	0.145	0.33	**8.39E-06**	2.55	
40	cg16933762	3:77147109	ROBO2	-	**2.77E-05**	10.36	**7.41E-05**	0.493	-1.12	**2.80E-05**	-10.61	
41	cg18282456	14:51561492	TRIM9	TRIM9	**2.85E-05**	-2.13	**7.52E-05**	0.430	-0.23	**7.62E-04**	1.83	
42	cg11530914	16:67281528	FHOD1;SLC9A5	FHOD1	**2.86E-05**	-1.87	**7.39E-05**	0.642	0.12	**7.74E-06**	1.95	
43	cg23414228	1:155043413	-	EFNA4	**2.96E-05**	-1.62	**7.36E-05**	0.627	-0.12	**1.11E-04**	1.48	
44	cg13323756	7:29847108	-	WIPF3	**2.97E-05**	3.90	**1.10E-05**	0.398	0.88	**0.015**	-2.39	✓
45	cg15268622	1:167189846	POU2F1	POU2F1	**3.00E-05**	-1.13	**7.54E-05**	0.247	-0.15	**2.50E-04**	1.00	
46	cg09811464	9:135465645	-	BARHL1	**3.04E-05**	-2.86	**3.49E-05**	0.168	-0.83	**0.008**	1.80	
47	cg24898753	11:61735917	FTH1	FTH1	**3.17E-05**	-2.79	**8.47E-05**	0.979	-0.02	**6.51E-04**	2.81	✓
48	cg08240832	1:241791729	OPN3	CHML	**3.21E-05**	7.50	**8.12E-05**	0.129	-1.81	**7.35E-06**	-10.09	✓
49	cg17998333	10:121356721	TIAL1	TIAL1	**3.30E-05**	-1.64	**7.53E-05**	0.441	-0.15	**6.05E-04**	1.41	
50	cg07170170	12:125672141	-	AACS	**3.37E-05**	-1.63	**7.21E-05**	0.986	0.00	**4.10E-04**	1.59	
51	cg15474615	3:134370182	KY	KY	**3.47E-05**	-2.93	**7.09E-05**	**0.049**	-1.43	0.055	1.39	
52	cg01492649	11:86749237	TMEM135	TMEM135	**3.68E-05**	-1.66	**7.65E-05**	0.965	0.01	**4.63E-05**	1.63	
53	cg09409821	7:98923250	ARPC1A	ARPC1A	**3.70E-05**	-1.62	**8.07E-05**	0.839	0.05	**4.47E-05**	1.72	
54	cg04108502	12:56360908	CDK2;SILV	CDK2	**3.71E-05**	-2.14	**9.64E-05**	0.200	-0.59	**0.007**	1.54	
55	cg27175943	16:85334190	-	KIAA0513	**3.72E-05**	-13.55	**9.95E-05**	0.108	-4.44	**0.034**	8.54	✓
56	cg25267526	19:12777743	MAN2B1;MORG1	WDR83	**3.74E-05**	-1.63	**9.77E-05**	0.943	-0.02	**5.29E-05**	1.56	
57	cg26609202	1:10532686	DFFA	DFFA	**3.74E-05**	-1.49	**7.93E-05**	0.423	-0.11	**2.12E-04**	1.39	
58	cg07622815	8:22526588	BIN3	BIN3	**4.00E-05**	-1.40	**8.16E-05**	0.579	0.12	**5.37E-05**	1.46	
59	cg23044270	3:43659607	ANO10	ANO10	**4.12E-05**	6.96	**1.96E-05**	0.410	0.98	**9.30E-04**	-6.72	✓
60	cg14572728	9:98279112	PTCH1	PTCH1	**4.16E-05**	-1.62	**1.02E-04**	0.561	-0.13	**3.92E-04**	1.46	
61	cg24701270	3:42947990	ZNF662	ZNF662	**4.19E-05**	-2.31	**1.08E-04**	0.400	-0.28	**1.38E-04**	1.96	
62	cg21196581	3:169780131	GPR160	GPR160	**4.19E-05**	7.09	**5.80E-05**	0.694	0.63	**0.010**	-6.15	
63	cg07192821	15:45003560	B2M	B2M	**4.21E-05**	-1.56	**7.84E-05**	0.275	-0.21	**1.14E-04**	1.34	
64	cg12207532	11:8040463	-	TUB	**4.28E-05**	-1.57	**1.14E-04**	0.995	0.00	**1.32E-04**	1.55	
65	cg24790706	7:157335647	PTPRN2	DNAJB6	**4.30E-05**	7.94	**1.13E-04**	0.240	1.52	**0.008**	-5.90	✓
66	cg09903872	7:157129451	DNAJB6	DNAJB6	**4.45E-05**	-1.48	**1.18E-04**	0.457	-0.16	**5.38E-04**	1.32	
67	cg07264124	19:11998686	ZNF69	ZNF69	**4.45E-05**	-1.65	**6.91E-05**	0.293	-0.26	**0.001**	1.39	
68	cg23436918	7:16685595	BZW2;ANKMY2	ANKMY2	**4.57E-05**	-1.46	**9.85E-05**	0.541	0.10	**1.74E-05**	1.56	
69	cg18004316	2:208634121	FZD5	FZD5	**4.63E-05**	-1.22	**1.05E-04**	0.646	0.07	**1.61E-04**	1.28	
70	cg12776966	19:30021276	VSTM2B	VSTM2B	**4.68E-05**	-2.05	**6.77E-05**	0.695	0.17	**0.002**	2.22	
71	cg06673826	6:111580383	KIAA1919	KIAA1919	**4.76E-05**	-1.13	**1.08E-04**	0.514	0.13	**1.32E-04**	1.26	
72	cg03958979	6:108486387	NR2E1	NR2E1	**4.77E-05**	-3.07	**1.13E-04**	0.144	-0.75	**0.009**	2.24	
73	cg09587503	5:134094082	DDX46	DDX46	**4.79E-05**	-1.29	**7.08E-05**	**0.025**	0.41	**2.62E-05**	1.64	

Probe information					5mC (OxBS)			Total modifications (BS)		5hmC (BS - OxBS)		
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	*P*	Δ	*P*	Δ	5hmC Positive probe
74	cg25463831	14:58766031	FLJ31306;ARID4A	ARID4A	**4.81E-05**	-1.05	**1.22E-04**	0.200	-0.16	**9.24E-04**	0.90	
75	cg27460531	17:78791723	RPTOR	CHMP6	**4.92E-05**	-5.48	**9.78E-05**	0.811	0.20	**3.15E-04**	5.53	✓

Supplementary Table 4: Significant neuropathology-associated differentially hydroxymethylated positions (DHPs) in the EC. Shown for each probe are chromosomal location (hg19), UCSC annotation, GREAT annotation, with corrected effect size (difference (Δ) in 5hmC levels) and corresponding P value between Braak score 0 and Braak score VI after adjusting for the covariates of age and sex, and an additional corresponding P value correcting for neuron/glia proportions (P CETS corrected), and annotation as to whether the mean 5hmC level was >0. Shown for all probes are Δ and corresponding P value for total DNA modifications and 5mC levels alone. All P values <0.05 are shown in bold.

Probe Information					5hmC (BS - OxBS)				Total modifications (BS)		5mC (OxBS)	
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	5hmC Positive probe	*P*	Δ	*P*	Δ
1	cg15147060	3:88108213	CGGBP1	CGGBP1	**3.94E-07**	1.73	**1.07E-06**		0.440	-0.13	**3.65E-09**	-1.92
2	cg25329939	13:98085947	RAP2A	RAP2A	**2.05E-06**	1.56	**4.93E-06**		0.074	0.31	**7.73E-06**	-1.29
3	cg17026303	3:38495161	LOC100128640;ACVR2B	ACVR2B	**2.74E-06**	2.39	**7.50E-06**		0.948	-0.02	**6.70E-07**	-2.46
4	cg00106685	3:52720133	GNL3;PBRM1	GNL3	**4.04E-06**	1.39	**6.63E-06**		0.413	-0.13	**7.80E-08**	-1.45
5	cg01710791	19:19144494	SFRS14;ARMC6	SUGP2	**5.07E-06**	1.41	**1.23E-05**		**0.024**	0.38	**3.03E-04**	-1.05
6	cg03160445	3:100120349	LNP1;TOMM70A	TOMM70A	**5.35E-06**	1.50	**1.41E-05**		0.680	0.07	**2.46E-05**	-1.48
7	cg05052335	5:130970657	RAPGEF6	RAPGEF6	**6.00E-06**	1.39	**1.48E-05**		0.215	-0.17	**9.73E-07**	-1.56
8	cg18070033	2:220025364	NHEJ1	NHEJ1	**6.24E-06**	1.84	**1.38E-05**		0.919	0.02	**3.64E-06**	-1.84
9	cg14252850	1:113010353	WNT2B	WNT2B	**6.26E-06**	-9.36	**1.30E-05**	✓	**0.004**	-4.68	**0.009**	4.82
10	cg08240832	1:241791729	OPN3	CHML	**7.35E-06**	-10.09	**1.99E-05**	✓	0.129	-1.81	**3.21E-05**	7.50
11	cg07536920	9:77112326	RORB	RORB	**7.66E-06**	1.91	**2.04E-05**		0.655	0.09	**5.99E-05**	-1.67
12	cg11530914	16:67281528	FHOD1;SLC9A5	FHOD1	**7.74E-06**	1.95	**2.09E-05**		0.642	0.12	**2.86E-05**	-1.87
13	cg12831866	20:24452036	TMEM90B	SYNDIG1	**8.16E-06**	2.44	**1.54E-05**		0.060	0.69	**0.004**	-1.66
14	cg13876163	1:45140495	C1orf228;TMEM53	TMEM53	**8.39E-06**	2.55	**1.45E-05**		0.145	0.33	**2.73E-05**	-2.25
15	cg03503087	10:118032626	GFRA1	GFRA1	**1.02E-05**	1.97	**1.79E-05**		0.731	0.06	**2.38E-06**	-1.89
16	cg07919466	19:50833956	KCNC3	KCNC3	**1.03E-05**	1.69	**1.68E-05**		0.730	-0.09	**8.41E-07**	-1.83
17	cg15481603	15:51977697	SCG3	SCG3	**1.26E-05**	-9.91	**2.98E-05**	✓	0.075	-2.75	**0.003**	6.16
18	cg03263685	2:68480160	PPP3R1	PPP3R1	**1.28E-05**	1.21	**3.25E-05**		0.932	0.01	**1.17E-05**	-1.23
19	cg12754733	19:11485436	C19orf39	SWSAP1	**1.29E-05**	1.78	**3.41E-05**		0.409	0.16	**5.61E-05**	-1.60
20	cg08888916	11:118016368	SCN4B	SCN4B	**1.50E-05**	1.72	**3.62E-05**		0.841	0.04	**5.44E-05**	-1.62
21	cg12144476	21:27541894	APP	APP	**1.61E-05**	1.44	**4.12E-05**		0.659	0.08	**4.85E-06**	-1.39
22	cg19370451	1:152386561	CRNN	CRNN	**1.63E-05**	-5.78	**3.42E-05**	✓	**0.002**	-3.04	**0.029**	2.73
23	cg23436918	7:16685595	BZW2;ANKMY2	ANKMY2	**1.74E-05**	1.56	**4.52E-05**		0.541	0.10	**4.57E-05**	-1.46
24	cg06174962	10:44351762	-	ZNF32	**1.88E-05**	-8.83	**4.81E-05**	✓	**0.024**	-2.39	**2.72E-04**	6.43
25	cg19005236	8:1571105	DLGAP2	DLGAP2	**1.93E-05**	-10.74	**2.48E-05**		0.341	-1.82	**2.68E-04**	8.76
26	cg16498314	16:30366841	CD2BP2	CD2BP2	**1.99E-05**	1.36	**3.81E-05**		0.271	0.18	**2.37E-04**	-1.11
27	cg05341384	16:86331661	-	FOXF1	**2.01E-05**	4.41	**4.84E-05**		0.875	-0.10	**4.15E-06**	-4.92
28	cg23665824	2:80530701	CTNNA2;LRRTM1	LRRTM1	**2.10E-05**	1.11	**4.98E-05**	✓	0.054	0.39	**6.15E-04**	-0.68
29	cg04149978	1:10764896	CASZ1	CASZ1	**2.17E-05**	-5.77	**2.28E-05**		0.310	-1.27	**0.003**	4.46
30	cg08318505	11:133903913	LOC100128239	JAM3	**2.26E-05**	2.20	**4.00E-05**		**0.001**	0.83	**0.003**	-1.31
31	cg10663655	7:66309627	LOC729156	TMEM248	**2.27E-05**	1.77	**5.56E-05**		0.061	0.28	**2.45E-04**	-1.49
32	cg02493644	7:84816314	-	SEMA3D	**2.43E-05**	1.60	**5.47E-05**		0.073	0.36	**3.34E-04**	-1.22
33	cg09587503	5:134094082	DDX46	DDX46	**2.62E-05**	1.64	**5.31E-05**		**0.025**	0.41	**4.79E-05**	-1.29

Probe Information					5hmC (BS - OxBS)				Total modifications (BS)		5mC (OxBS)	
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	5hmC Positive probe	*P*	Δ	*P*	Δ
34	cg08182454	19:46529189	-	PGLYRP1	**2.67E-05**	-10.76	**6.58E-05**	✓	0.106	-2.73	**0.001**	7.40
35	cg20066612	1:160050948	KCNJ9	KCNJ9	**2.70E-05**	5.84	**6.47E-05**	✓	**6.79E-06**	3.25	**0.027**	-2.89
36	cg13009654	5:137802252	EGR1	EGR1	**2.71E-05**	3.94	**3.00E-05**		0.070	1.01	**8.30E-04**	-2.91
37	cg16933762	3:77147109	ROBO2	-	**2.80E-05**	-10.61	**7.11E-05**		0.493	-1.12	**2.77E-05**	10.36
38	cg25954627	10:8090846	-	GATA3	**3.31E-05**	3.38	**7.17E-05**		0.484	0.29	**3.39E-04**	-3.00
39	cg18494448	17:16492891	-	ZNF287	**3.41E-05**	1.59	**9.07E-05**		0.150	0.26	**1.75E-04**	-1.40
40	cg05503627	12:26267490	-	BHLHE41	**3.47E-05**	0.88	**4.10E-05**	✓	**0.002**	0.54	0.197	-0.23
41	cg20894246	22:41682263	RANGAP1	RANGAP1	**3.48E-05**	2.24	**9.19E-05**		0.108	0.38	**4.97E-04**	-1.86
42	cg03436461	17:13504687	HS3ST3A1	HS3ST3A1	**3.68E-05**	2.08	**9.77E-05**		0.565	0.17	**1.39E-04**	-1.91
43	cg17337106	8:80577176	STMN2	STMN2	**3.80E-05**	-6.70	**9.62E-05**		0.075	-2.19	**8.36E-04**	3.93
44	cg17848546	1:209605304	LOC642587;MIR205	CAMK1G	**3.82E-05**	-7.82	**9.14E-05**	✓	**0.035**	-2.68	**0.007**	4.61
45	cg09504196	7:148823235	ZNF425;ZNF398	ZNF425	**3.85E-05**	1.58	**7.87E-05**		0.113	0.29	**1.54E-04**	-1.26
46	cg15624314	1:226251003	H3F3A;LOC440926	H3F3A	**4.30E-05**	1.59	**9.50E-05**		0.320	-0.25	**6.14E-07**	-1.84
47	cg09409821	7:98923250	ARPC1A	ARPC1A	**4.47E-05**	1.72	**9.14E-05**		0.839	0.05	**3.70E-05**	-1.62
48	cg19878762	22:23484297	RTDR1	RTDR1	**4.49E-05**	1.49	**1.08E-04**		0.126	0.28	**4.17E-04**	-1.24
49	cg22026192	12:124944377	NCOR2	NCOR2	**4.51E-05**	-8.41	**1.14E-04**	✓	**0.013**	-2.91	**0.003**	5.56
50	cg01492649	11:86749237	TMEM135	TMEM135	**4.63E-05**	1.63	**9.62E-05**		0.965	0.01	**3.68E-05**	-1.66
51	cg20966551	19:12949060	MAST1	MAST1	**4.66E-05**	1.59	**1.02E-04**		0.231	0.23	**9.07E-05**	-1.39
52	cg14546505	9:102863959	INVS	INVS	**4.69E-05**	-7.18	**6.03E-05**	✓	0.089	-2.41	**0.020**	4.28
53	cg17158913	1:10764886	CASZ1	CASZ1	**4.78E-05**	-8.74	**9.33E-05**		0.366	-1.16	**8.94E-04**	7.56
54	cg14717061	11:126138971	FOXRED1;SRPR	FOXRED1	**4.89E-05**	1.28	**7.78E-05**		0.562	0.08	**7.13E-05**	-1.20
55	cg07379167	10:131769074	-	EBF3	**4.89E-05**	2.45	**1.31E-04**		0.062	-0.77	**5.45E-05**	-3.02
56	cg07269003	3:142681516	PAQR9	PAQR9	**4.90E-05**	-8.72	**6.60E-05**		**0.021**	-3.64	**0.009**	4.74
57	cg24432048	1:3827651	LOC100133612	C1orf174	**4.97E-05**	1.91	**6.87E-05**		**0.050**	0.55	**0.001**	-1.29

Supplementary Table 5: Significant neuropathology-associated differentially methylated positions (DMPs) in the CER. Shown for each probe are chromosomal location (hg19), UCSC annotation, GREAT annotation, with corrected effect size (difference (Δ) in 5mC levels) and corresponding P value between Braak score 0 and Braak score VI after adjusting for the covariates of age and sex, and an additional corresponding P value also correcting for neuron/glia proportions (P CETS corrected). Shown for all probes are Δ and corresponding P value for total DNA modifications and 5hmC levels alone, with annotation as to whether the mean 5hmC level was >0. All P values <0.05 are shown in bold.

Probe information					**5mC (OxBS)**			**Total modifications (BS)**		**5hmC (BS - OxBS)**		
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	*P*	Δ	*P*	Δ	5hmC Positive probe
1	cg04868540	11:134263237	B3GAT1	B3GAT1	**7.05E-07**	7.69	8.14E-07	8.31E-01	0.23	3.99E-05	-7.42	✓
2	cg23448978	7:51209365	COBL	COBL	**5.10E-06**	-1.67	4.45E-06	2.32E-02	-0.58	1.11E-02	1.09	✓
3	cg12898370	12:24737779	C12orf67	BCAT1	**1.13E-05**	5.98	4.93E-06	6.04E-01	-0.91	1.85E-03	-6.78	✓
4	cg21548096	12:69326897	CPM	CPM	**2.21E-05**	0.80	2.48E-05	5.37E-01	0.11	5.96E-03	-0.68	✓
5	cg08156809	7:74071310	GTF2I	GTF2I	**3.23E-05**	5.78	3.10E-05	4.89E-01	-0.55	3.32E-05	-6.39	
6	cg21858113	11:118015340	SCN4B	SCN4B	**4.09E-05**	5.05	2.14E-05	2.20E-01	1.22	1.84E-04	-3.88	✓
7	cg02592271	2:27665507	KRTCAP3	KRTCAP3	**4.12E-05**	-9.06	1.96E-05	6.36E-02	-3.12	1.54E-02	5.87	✓

Supplementary Table 6: Significant neuropathology-associated differentially hydroxymethylated positions (DHPs) in the CER. Shown for each probe are chromosomal location (hg19), UCSC annotation, GREAT annotation, with corrected effect size (difference (Δ) in 5hmC levels) and corresponding P value between Braak score 0 and Braak score VI after adjusting for the covariates of age and sex, and an additional corresponding P value correcting for neuron/glia proportions (P CETS corrected), and annotation as to whether the mean 5hmC level was >0. Shown for all probes are Δ and corresponding P value for total DNA modifications and 5mC levels alone. All P values <0.05 are shown in bold.

Probe Information					5hmC (BS - OxBS)				Total modifications (BS)		5mC (OxBS)	
Rank	Probe	Position	UCSC Gene Annotation	GREAT Annotation	*P*	Δ	*P* CETS corrected	5hmC Positive probe	*P*	Δ	*P*	Δ
1	cg09214953	12:111800702	FAM109A	FAM109A	**4.40E-06**	-5.70	6.26E-06	✓	3.31E-03	-2.40	4.51E-03	3.33
2	cg01385157	15:89421347	HAPLN3	HAPLN3	**1.35E-05**	-8.82	1.89E-05	✓	1.72E-01	-5.63	3.39E-01	3.40
3	cg04116354	1:26003643	MAN1C1	MAN1C1	**1.41E-05**	10.67	1.72E-05	✓	1.12E-02	3.96	1.55E-03	-6.71
4	cg10231603	1:155829206	SYT11	GON4L	**2.11E-05**	-2.75	1.39E-05		1.78E-01	-0.59	1.36E-03	2.17
5	cg14289985	19:57018848	ZNF471	ZNF471	**2.61E-05**	1.53	3.16E-05	✓	4.17E-05	1.07	1.61E-01	-0.44
6	cg17606548	5:66123702	MAST4	MAST4	**2.66E-05**	-8.44	2.89E-05	✓	1.76E-01	-1.87	2.20E-04	6.70
7	cg16061228	11:118014547	SCN4B	SCN4B	**2.89E-05**	-5.40	4.03E-05	✓	2.50E-01	-1.22	1.29E-02	4.21
8	cg08156809	7:74071310	GTF2I	GTF2I	**3.32E-05**	-6.39	2.43E-05		4.89E-01	-0.55	3.23E-05	5.78
9	cg23180780	2:12260486	-	LPIN1	**3.39E-05**	-5.74	4.16E-05		5.16E-02	-1.94	9.08E-03	3.83
10	cg00732538	16:87635674	JPH3	JPH3	**3.41E-05**	-2.78	3.25E-05		5.08E-01	-0.19	6.53E-05	2.57
11	cg10461218	12:124144432	GTF2H3	TCTN2	**3.85E-05**	-5.54	3.81E-05	✓	1.59E-01	-1.20	9.88E-05	4.33
12	cg04868540	11:134263237	B3GAT1	B3GAT1	**3.99E-05**	-7.42	5.22E-05	✓	8.31E-01	0.23	7.05E-07	7.69

Supplementary Table 7: Pathways enriched with neuropathology-associated DMPs in EC. Abbreviations: Gene Ontology (GO), Biological Process (BP), Cellular Component (CC), Molecular Function (MF). Shown are pathways were $P < 0.005$. Data is organized with the most significant pathways at the top of the list.

GO Term	GO Name	Type	No. Genes in Pathway	No. Test List in Pathway	*P* Genes in Test List
GO:0060715	syncytiotrophoblast cell differentiation involved in labyrinthine layer development	BP	3	2	8.26E-05
GO:0016199	axon midline choice point recognition	BP	4	2	2.30E-04
GO:0060717	chorion development	BP	7	2	3.92E-04
GO:0000805	X chromosome	CC	7	2	4.15E-04
GO:0016198	axon choice point recognition	BP	6	2	4.20E-04
GO:0005773	vacuole	CC	1157	13	5.44E-04
GO:1903867	extraembryonic membrane development	BP	9	2	6.07E-04
GO:0021772	olfactory bulb development	BP	32	3	6.14E-04
GO:0021988	olfactory lobe development	BP	33	3	6.86E-04
GO:0042246	tissue regeneration	BP	50	3	8.39E-04
GO:0002726	positive regulation of T cell cytokine production	BP	14	2	9.35E-04
GO:0044440	endosomal part	CC	416	7	0.001
GO:0044437	vacuolar part	CC	684	9	0.001
GO:0002724	regulation of T cell cytokine production	BP	19	2	0.001
GO:0045893	positive regulation of transcription, DNA-templated	BP	1301	14	0.001
GO:1903508	positive regulation of nucleic acid-templated transcription	BP	1301	14	0.001
GO:1902680	positive regulation of RNA biosynthetic process	BP	1322	14	0.002
GO:0021889	olfactory bulb interneuron differentiation	BP	12	2	0.002
GO:0051254	positive regulation of RNA metabolic process	BP	1367	14	0.002
GO:0038023	signaling receptor activity	MF	1227	10	0.002
GO:0030033	microvillus assembly	BP	16	2	0.002
GO:0009880	embryonic pattern specification	BP	58	3	0.003
GO:0031904	endosome lumen	CC	22	2	0.003
GO:0030139	endocytic vesicle	CC	246	5	0.003
GO:0002369	T cell cytokine production	BP	25	2	0.003
GO:0034219	carbohydrate transmembrane transport	BP	26	2	0.003
GO:0035994	response to muscle stretch	BP	19	2	0.003
GO:0043235	receptor complex	CC	318	6	0.003
GO:0004888	transmembrane signaling receptor activity	MF	1119	9	0.003
GO:0030666	endocytic vesicle membrane	CC	147	4	0.003
GO:0060644	mammary gland epithelial cell differentiation	BP	16	2	0.003
GO:0033267	axon part	CC	206	5	0.003
GO:0030673	axolemma	CC	14	2	0.003
GO:0006357	regulation of transcription from RNA polymerase II promoter	BP	1761	16	0.003

GO Term	GO Name	Type	No. Genes in Pathway	No. Test List in Pathway	*P* Genes in Test List
GO:0006366	transcription from RNA polymerase II promoter	BP	1947	17	0.003
GO:0010628	positive regulation of gene expression	BP	1616	15	0.003
GO:0008093	cytoskeletal adaptor activity	MF	15	2	0.003
GO:0044304	main axon	CC	56	3	0.003
GO:0005768	endosome	CC	771	9	0.004
GO:0019731	antibacterial humoral response	BP	40	2	0.004
GO:0060713	labyrinthine layer morphogenesis	BP	21	2	0.004
GO:0021872	forebrain generation of neurons	BP	63	3	0.004
GO:0007176	regulation of epidermal growth factor-activated receptor activity	BP	25	2	0.004
GO:0098659	inorganic cation import into cell	BP	25	2	0.004
GO:0099587	inorganic ion import into cell	BP	25	2	0.004
GO:0010008	endosome membrane	CC	386	6	0.004
GO:0002711	positive regulation of T cell mediated immunity	BP	31	2	0.004
GO:0002720	positive regulation of cytokine production involved in immune response	BP	31	2	0.005
GO:0099600	transmembrane receptor activity	MF	1173	9	0.005
GO:0030332	cyclin binding	MF	20	2	0.005
GO:0019730	antimicrobial humoral response	BP	45	2	0.005
GO:0000803	sex chromosome	CC	27	2	0.005
GO:0032528	microvillus organization	BP	24	2	0.005

Supplementary Table 8: Pathways enriched with neuropathology-associated DHPs in EC. Abbreviations: Gene Ontology (GO), Biological Process (BP), Cellular Component (CC), Molecular Function (MF). Shown are pathways were $P < 0.005$. Data is organized with the most significant pathways at the top of the list.Data is organized with the most significant pathways at the top of the list.

GO Term	GO Name	Type	No. Genes in Pathway	No. Test List in Pathway	*P* Genes in Test List
GO:0050803	regulation of synapse structure or activity	BP	115	5	8.24E-05
GO:0016199	axon midline choice point recognition	BP	4	2	1.30E-04
GO:0016198	axon choice point recognition	BP	6	2	2.55E-04
GO:0050807	regulation of synapse organization	BP	111	4	9.11E-04
GO:0050808	synapse organization	BP	229	5	0.001
GO:0030033	microvillus assembly	BP	16	2	0.002
GO:0007416	synapse assembly	BP	135	4	0.002
GO:0033233	regulation of protein sumoylation	BP	20	2	0.002
GO:0030424	axon	CC	398	6	0.003
GO:0032528	microvillus organization	BP	24	2	0.003
GO:0051963	regulation of synapse assembly	BP	78	3	0.004

Supplementary Table 9: Table highlighting specific genomic features enriched for neuropathology-associated DMPs. Taking the genomic location of neuropathology-associated DMPs above our significance threshold ($P < 5 \times 10^{-5}$), we used a Fisher's exact test to examine whether hypermethylated or hypomethylated loci are enriched within specific genomic features compared to all Braak-associated probes ($P < 0.05$). Shown for each Slieker genomic feature is whether we observed significant disease-associated hypermethylation or hypomethylation, direction of change and corresponding P value. Abbreviations: CpG island (CGI), outside CpG island (nonCGI or NC), CpG island shelf (SHE), CpG island shore (SHO), distal promoter (DP), downstream (DS), gene body (GB), intergenic (IG), proximal promoter (PP), unannotated (UA).

		Significant Hypermethylation	Direction	*P* value	Significant Hypomethylation	Direction	*P* value
DP	CGI	✓	-	7.01E-29	✓	+	2.08E-42
	NC				✓	-	1.14E-07
	SHE				✓	-	1.75E-21
	SHO	✓	-	7.61E-04			
DS	CGI	✓	-	1.50E-07	✓	+	0.008
	NC				✓	-	1.72E-06
	SHE				✓	-	0.003
	SHO				✓	-	0.001
GB	CGI	✓	-	3.65E-108	✓	+	1.56E-64
	NC	✓	+	1.08E-05	✓	-	2.55E-168
	SHE				✓	-	6.84E-41
	SHO	✓	-	1.04E-30	✓	+	3.43E-05
IG	CGI	✓	-	1.04E-28	✓	+	5.60E-22
	NC	✓	+	1.46E-04	✓	-	5.35E-84
	SHE				✓	-	3.92E-09
	SHO	✓	-	1.74E-04	✓	+	0.001
PP	CGI	✓	-	4.94E-324	✓	+	<1.00E-330
	NC	✓	+	0.012	✓	-	1.93E-29
	SHE	✓	-	0.004	✓	+	0.037
	SHO	✓	-	6.10E-128	✓	+	1.48E-49
	UA	✓	+	8.35E-08	✓	-	3.47E-73

Supplementary Table 10: Table highlighting specific genomic features enriched for neuropathology-associated DHPs. Taking the genomic location of neuropathology-associated DHPs above our significance threshold ($P < 5 \times 10^{-5}$), we used a Fisher's exact test to examine whether hyperhydroxymethylated or hypohydroxymethylated loci are enriched within specific genomic features compared to all Braak-associated probes ($P < 0.05$). Shown for each Slieker genomic feature is whether we observed significant disease-associated hyperhydroxymethylation or hypohydroxymethylation, direction of change and corresponding P value. Abbreviations: CpG island (CGI), outside CpG island (nonCGI or NC), CpG island shelf (SHE), CpG island shore (SHO), distal promoter (DP), downstream (DS), gene body (GB), intergenic (IG), proximal promoter (PP), unannotated (UA).

		Significant Hyperhydroxymethylation	Direction	*P* value	Significant Hypohydroxymethylation	Direction	*P* value
DP	CGI	✓	+	1.62E-10	✓	-	3.30E-33
	NC	✓	-	6.28E-19	✓	+	1.25E-05
	SHE	✓	-	4.54E-41			
	SHO	✓	+	2.86E-07	✓	-	1.16E-14
DS	CGI				✓	-	3.93E-10
	NC	✓	-	2.89E-14	✓	+	0.037
	SHE	✓	-	1.78E-06	✓	+	0.028
	SHO	✓	-	3.09E-09	✓	+	3.21E-04
GB	CGI	✓	+	9.97E-09	✓	-	1.15E-100
	NC	✓	-	4.94E-324	✓	+	1.69E-08
	SHE	✓	-	2.64E-111	✓	-	1.57E-05
	SHO	✓	+	7.09E-43	✓	-	3.43E-55
IG	CGI	✓	+	1.59E-05	✓	-	1.64E-46
	NC	✓	-	6.27E-165	✓	+	3.19E-09
	SHE	✓	-	3.52E-27			
	SHO	✓	+	4.17E-17	✓	-	8.78E-19
PP	CGI	✓	+	<1.00E-330	✓	-	4.94E-324
	NC	✓	-	7.26E-61	✓	+	2.10E-11
	SHE	✓	-	3.22E-08	✓	+	4.17E-05
	SHO	✓	+	3.76E-26	✓	-	1.29E-257
	UA	✓	-	2.83E-53	✓	+	7.22E-04

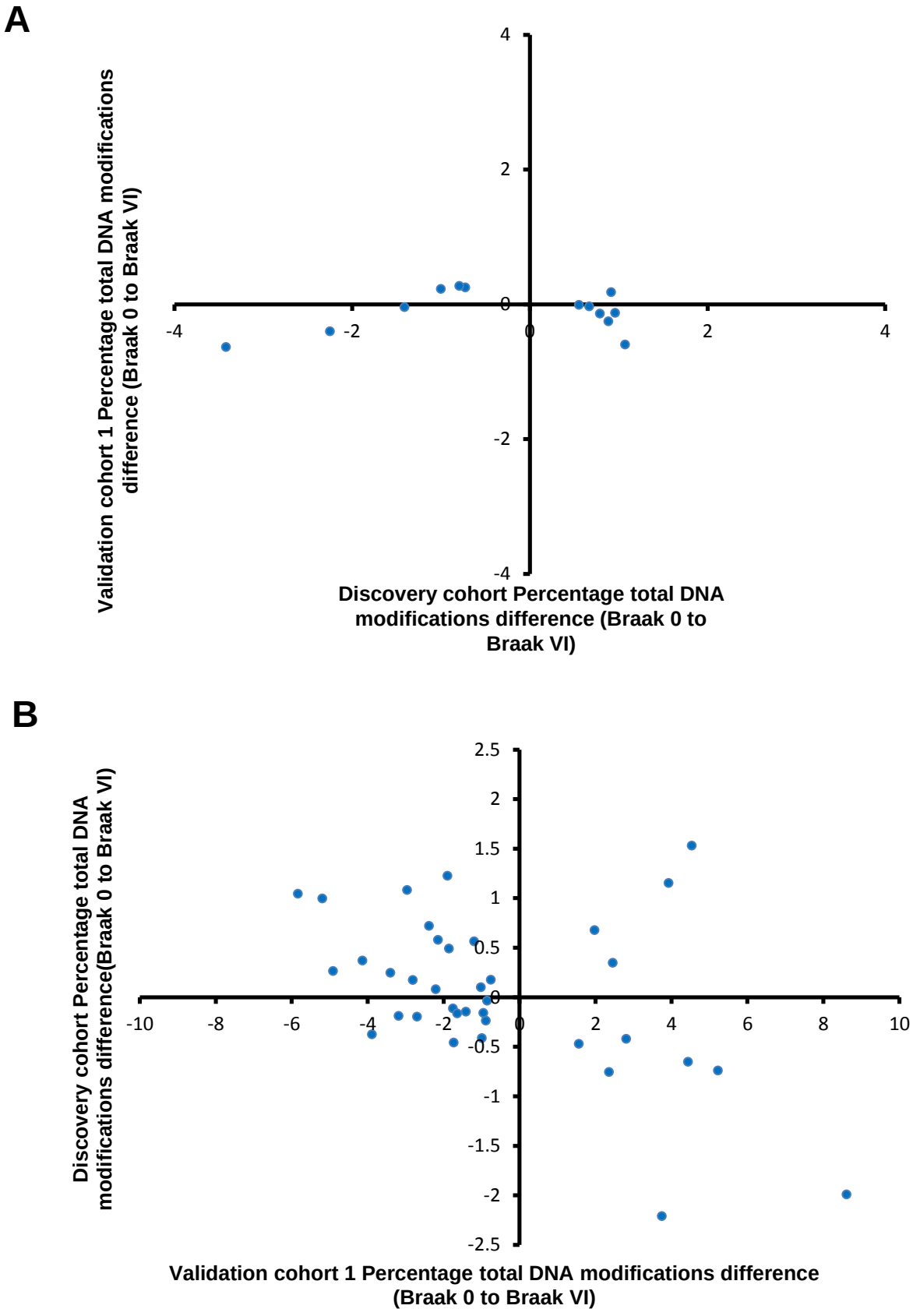
Supplementary Table 11: Sample and demographic information for the three cohorts used. (**A**) The discovery cohort consisted of 450K array BS and OxBS data generated in 96 individuals from the MRC London Brain Bank for Neurodegenerative Disease. (**B**) Validation cohort 1 consisted of previously published 450K array BS data generated in an independent cohort of 117 individuals also from the MRC London Brain Bank for Neurodegenerative Disease [7]. (**C**) Validation cohort 2 consisted of pyrosequencing BS and OxBS data we generated in an independent cohort of 96 individuals from the Thomas Willis Oxford Brain Collection. For the discovery cohort and validation cohort 1 we analyzed data from two brain regions: the entorhinal cortex and cerebellum. For validation cohort 2 we analyzed ANK1 pyrosequencing data from the entorhinal cortex only. Shown for each dataset are the number of samples (N) that passed QC , distribution of sex, clinical diagnosis (Alzheimer's disease (AD) or non-demented control (CTL)), Braak stage, Mean age and Postmortem interval (± standard deviation (SD)).

A		Entorhinal Cortex			Cerebellum		
		Total Modifications (BS)	5mC (OxBS)	5hmC (BS - OxBS)	Total Modifications (BS)	5mC (OxBS)	5hmC (BS - OxBS)
N passed QC		91	85	85	95	94	94
Gender [M/F] (%)		(56) 51/40 (44)	(56.5) 48/37 (43.5)	(56.5) 48/37 (43.5)	(56.8) 54/41 (43.2)	(57.4) 54/40 (42.6)	(57.4) 54/40 (42.6)
Mean age (±SD)		81.2 (9.5)	81.3 (9.5)	81.2 (9.5)	81.2 (9.3)	81.2 (9.3)	81.2 (9.3)
Diagnosis [AD/CTL] (%)		(70.3) 64/27 (29.7)	(70.6) 60/25 (29.4)	(70.6) 60/25 (29.4)	(70.5) 67/28 (29.5)	(70.2) 66/28 (29.8)	(70.2) 66/28 (29.8)
Braak Stage	0	8	7	7	8	8	8
	I	3	3	3	3	3	3
	II	11	10	10	12	12	12
	III	6	6	6	6	6	6
	IV	8	7	7	10	10	10
	V	18	17	17	19	18	18
	VI	37	35	35	37	37	37
Mean postmortem Interval [minutes] (±SD)		2539.5 (1288.1)	2490.7 (1288.5)	2490.7 (1288.5)	2576.5 (1315.2)	2581.6 (1321.3)	2581.6 (1321.3)

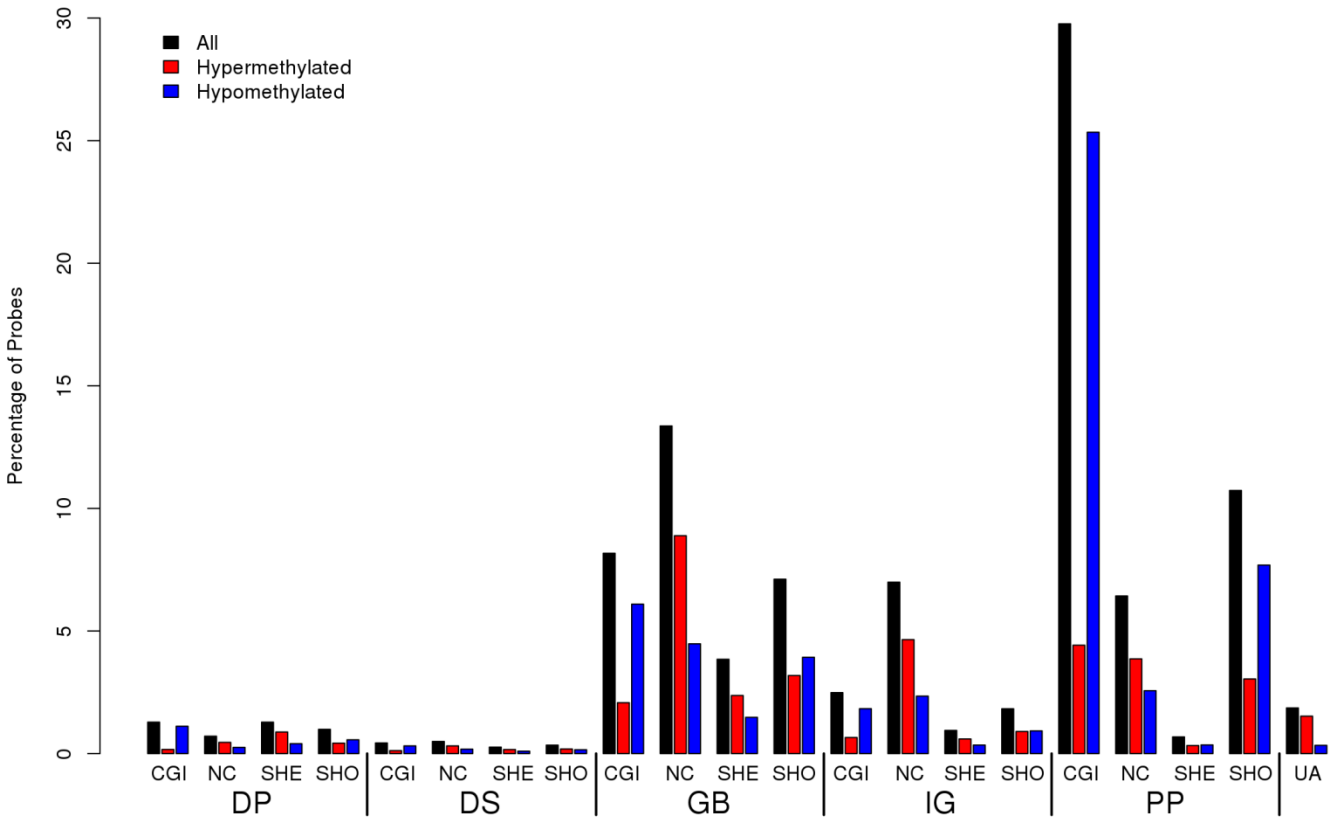
B		Entorhinal Cortex	Cerebellum
		Total Modifications (BS)	
N passed QC		104	108
Gender [M/F] (%)		(40.4) 42/62 (59.6)	(42.6) 46/62 (57.4)
Mean age (±SD)		84.9 (8.7)	83.8 (10.0)
Diagnosis [AD/CTL] (%)		(59.6) 62/42 (40.4)	(59.3) 64/44 (40.7)
Braak Stage	0	5	8
	I	11	11
	II	8	7
	III	13	13
	IV	5	5
	V	18	19
	VI	44	45
Mean postmortem Interval [minutes] (±SD)		1997.6 (1227.0)	2071.7 (1303.6)

C		Entorhinal Cortex		
		Total Modifications (BS)	5mC (OxBS)	5hmC (BS - OxBS)
N passed QC		96	92	92
Gender [M/F] (%)		(56.3) 54/42 (43.8)	(56.5) 52/40 (43.5)	(56.5) 52/40 (43.5)
Mean age (±SD)		85.0 (7.2)	84.8 (7.3)	84.8 (7.3)
Diagnosis [AD/CTL] (%)		(50) 48/48 (50)	(50) 46/46 (50)	(50) 46/46 (50)
Braak Stage	0	5	5	5
	I	6	6	6
	II	37	35	35
	III	0	0	0
	IV	0	0	0
	V	23	22	22
	VI	25	24	24
Mean postmortem Interval [minutes] (±SD)		2960.6 (1943.8)	2958.9 (1958.7)	2958.9 (1958.7)

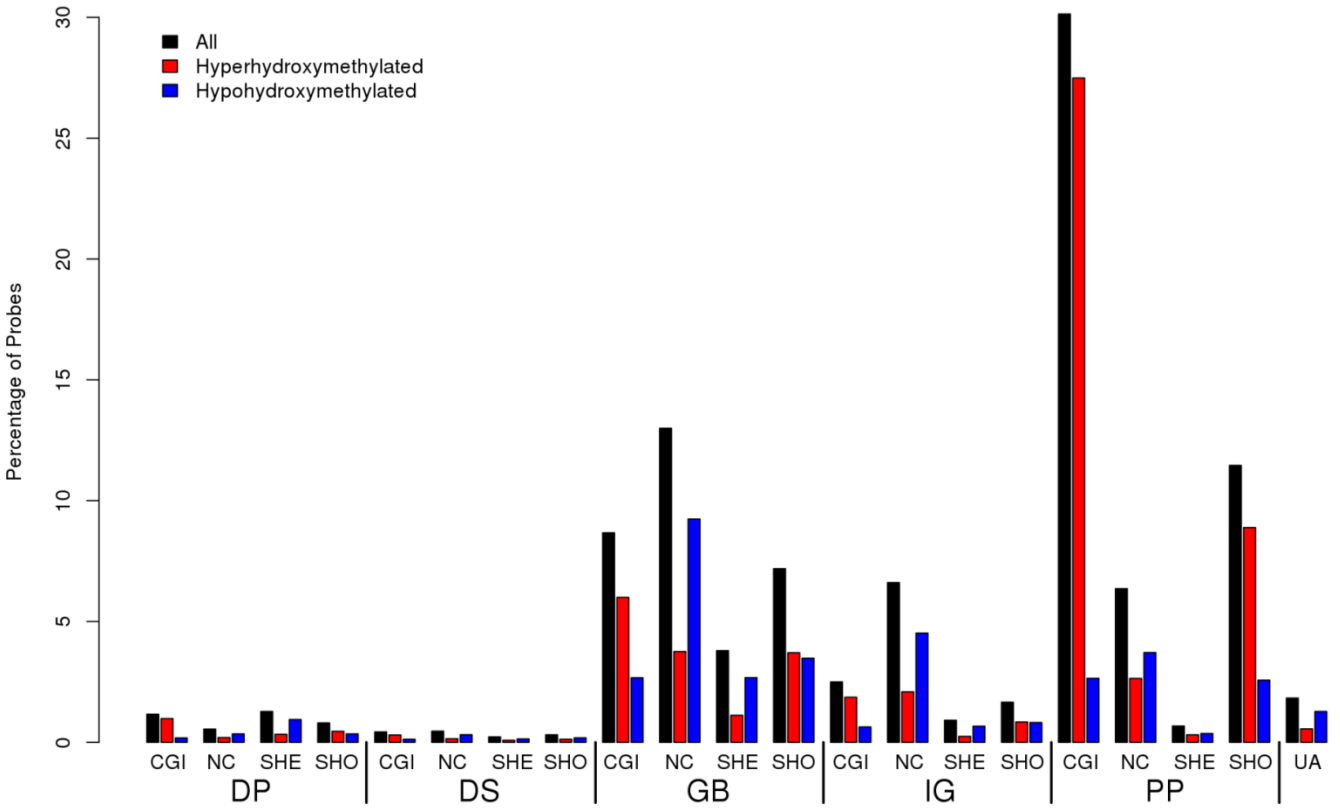
Supplementary Figure 1: Braak-associated DMoPs in the CER are not consistent between studies (**A**) There was no consistent direction of effect when comparing Braak-associated DMoPs ($P < 5 \times 10^{-5}$) in the CER in the discovery cohort with the effect size of the same probes in a previously published dataset of Braak-associated DMoPs in the CER [7] (validation cohort 1) (sign test $P = 0.267$). (**B**) There was also no consistent direction of effect for previously published Braak-associated DMoPs in the CER [7] ($P < 5 \times 10^{-5}$) (validation cohort 1) when compared with the effect size of the same probes in the current study (sign test $P = 0.188$).



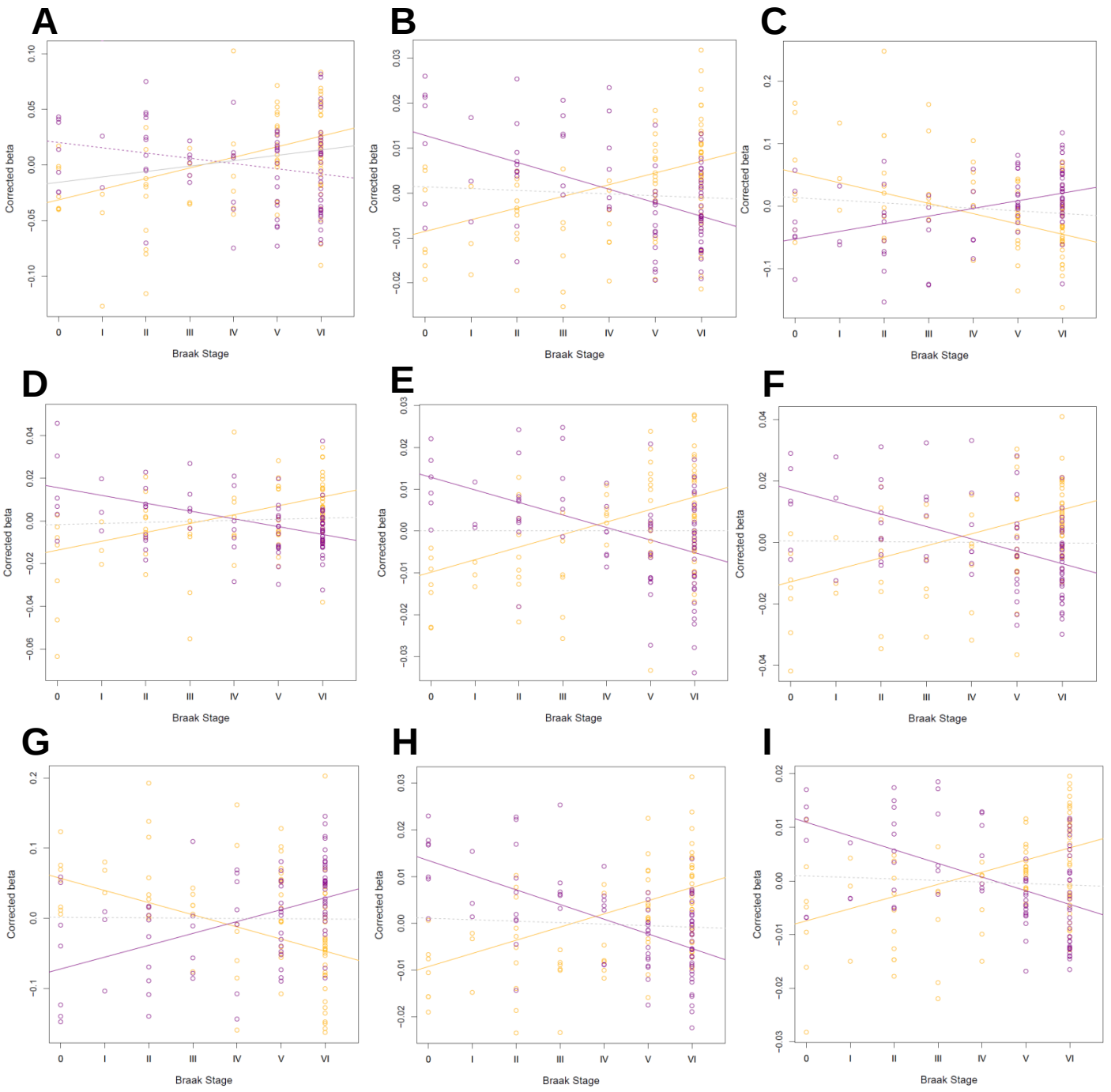
Supplementary Figure 2: The most significant neuropathology-associated DMPs are enriched at specific genomic features in the EC. Taking the genomic location of neuropathology-associated DMPs above our significance threshold ($P < 5 \times 10^{-5}$), we used a Fisher's exact test to examine whether these hypermethylated or hypomethylated loci are enriched within specific genomic features when compared to all Braak-associated DMPs ($P < 0.05$) with results of the enrichments tests shown in Supplementary Table 9. Abbreviations: CpG island (CGI), outside CpG island (nonCGI or NC), CpG island shelf (SHE), CpG island shore (SHO), distal promoter (DP), downstream (DS), gene body (GB), intergenic (IG), proximal promoter (PP), unannotated (UA).



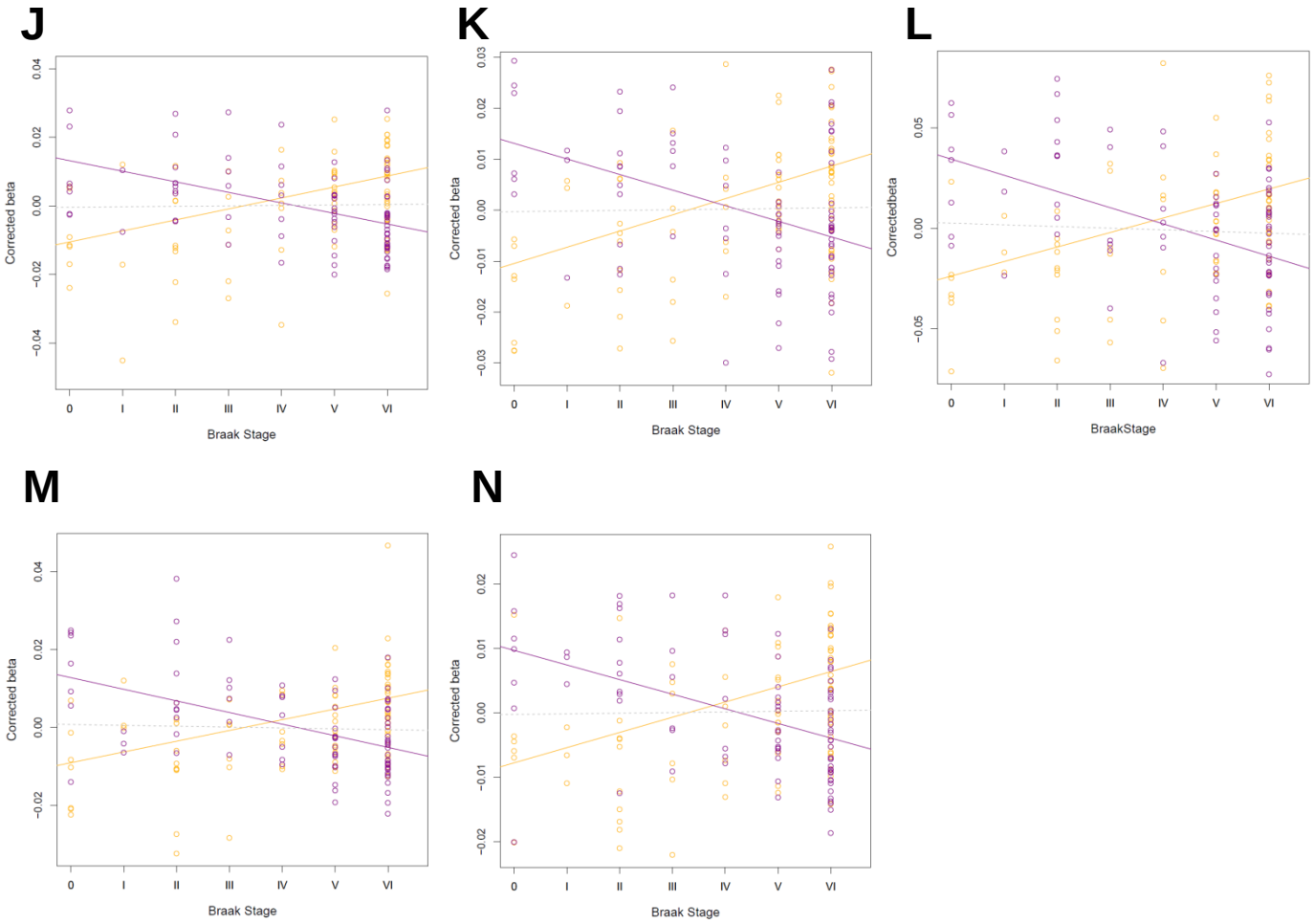
Supplementary Figure 3: The most significant neuropathology-associated DHPs are enriched at specific genomic features in the EC. Taking the genomic location of neuropathology-associated DHPs above our significance threshold ($P < 5 \times 10^{-5}$), we used a Fisher's exact test to examine whether these hyperhydroxymethylated or hypohydroxymethylated loci are enriched within specific genomic features when compared to all Braak-associated DHPs ($P < 0.05$) , with results of the enrichments tests shown in Supplementary Table 10. Abbreviations: CpG island (CGI), outside CpG island (nonCGI or NC), CpG island shelf (SHE), CpG island shore (SHO), distal promoter (DP), downstream (DS), gene body (GB), intergenic (IG), proximal promoter (PP), unannotated (UA).



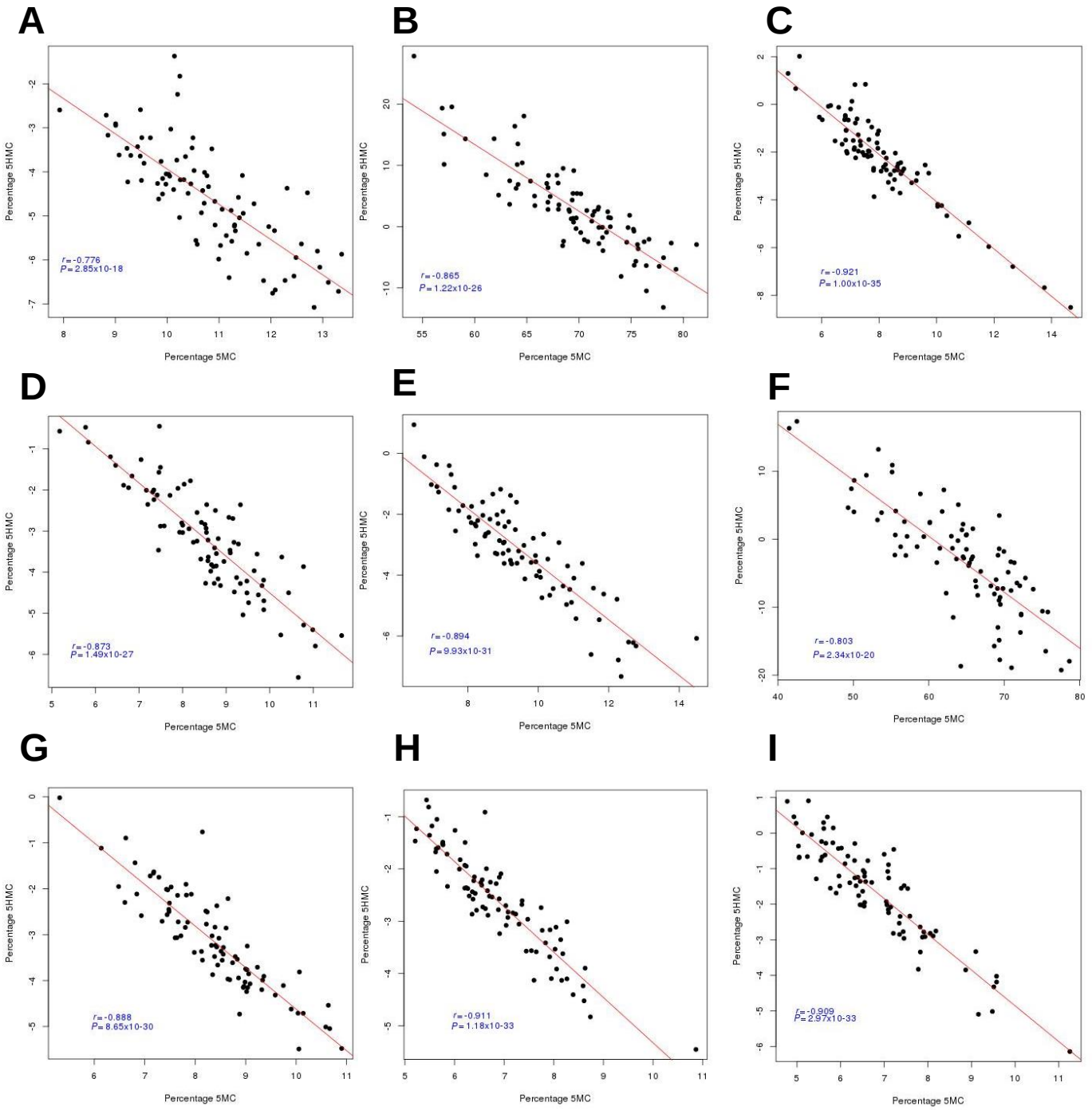
Supplementary Figure 4: Scatter plot of beta values against Braak stage for the 14 probes that feature in Table 1. Shown are beta values, corrected for age and gender, for (**A**) cg20066612 (*KCNJ9*), (**B**) cg15624314 (*H3F3A*), (**C**) cg08240832 (*OPN3*), (**D**) cg13876163 (*TMEM53*), (**E**) cg18070033 (*NHEJ1*), (**F**) cg17026303 (*ACVR2B*), (**G**) cg16933762 (*ROBO2*), (**H**) cg15147060 (*CGGBP1*), (**I**) cg05052335 (*RAPGEF6*), (**J**) cg03503087 (*GFRA1*), (**K**) cg11530914 (*FHOD1*), (**L**) cg05341384 (*FOXF1*), (**M**) cg07919466 (*KCNC3*) and (**N**) cg12144476 (*APP*). Shown are beta values corrected for age and gender (Y axis), against Braak stage (X axis). Purple circles represents 5mC levels (OxBS), yellow circles represents 5hmC levels (BS-OxBS). Regression lines are shown for 5mC (purple line), 5hmC (yellow line) and total DNA modifications (grey line). Solid regression lines indicate $P < 0.05$, whilst dashed lines indicate $P > 0.05$.



Supplementary Figure 4 (cont.)

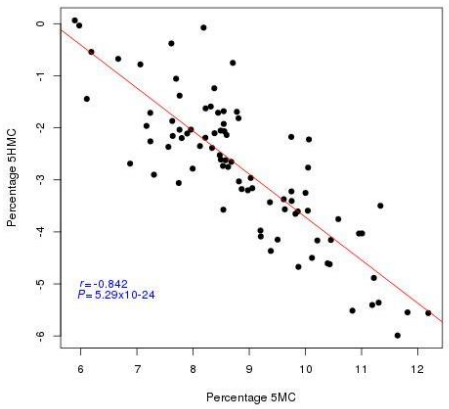


Supplementary Figure 5: Correlation plots of 5mC levels and 5hmC levels for probes that feature in Table 1 with no net change in total DNA modifications. Shown are beta levels, corrected for age and gender for (**A**) cg15624314 (*H3F3A*), (**B**) cg08240832 (*OPN3*), (**C**) cg13876163 (*TMEM53*), (**D**) cg18070033 (*NHEJ1*), (**E**) cg17026303 (*ACVR2B*), (**F**) cg16933762 (*ROBO2*), (**G**) cg15147060 (*CGGBP1*), (**H**) cg05052335 (*RAPGEF6*), (**I**) cg03503087 (*GFRA1*), (**J**) cg11530914 (*FHOD1*), (**K**) cg05341384 (*FOXF1*), (**L**) cg07919466 (*KCNC3*) and (**M**) cg12144476 (*APP*). Plotted is the % 5hmC (Y axis) against % 5mC after correcting for age and gender for all samples.

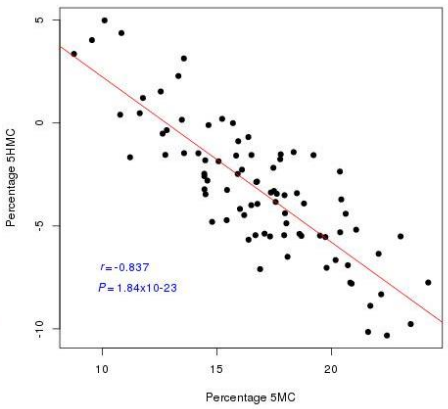


Supplementary Figure 5 (cont.)

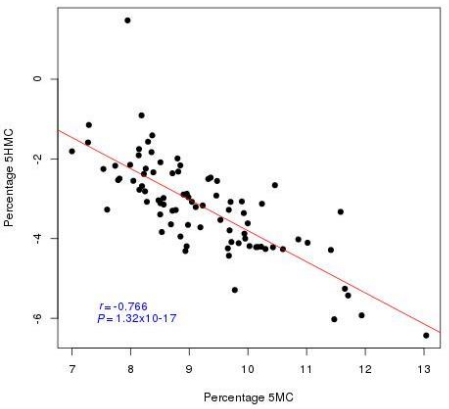
J



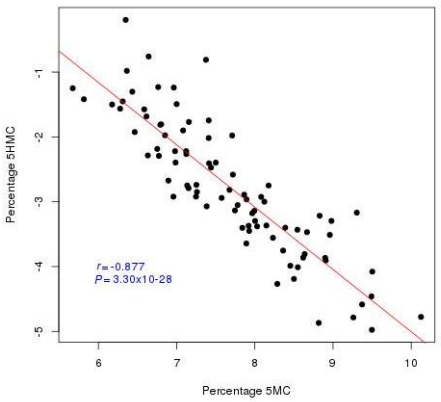
K



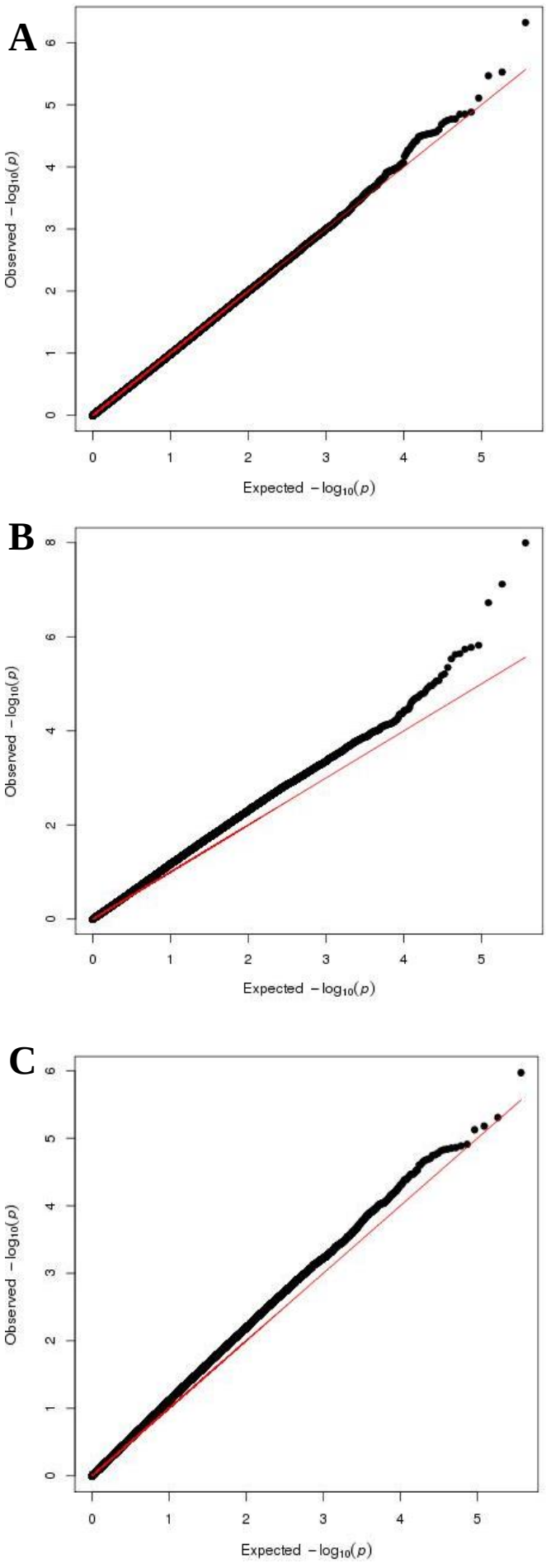
L



M



Supplementary Figure 6: Quantile-quantile (Q-Q) plots of expected versus observed *P* value to check for inflation in linear regression analyses. The Q-Q plots for (**A**) total DNA modifications (BS data), (**B**) DNA methylation (OxBS data) and (**C**) DNA hydroxmethylation (BS – OxBS data) in the EC shows no *P* value inflation in our analyses.



APPENDIX 4 – A CROSS-BRAIN-REGIONS STUDY OF ANK1 DNA METHYLATION IN DIFFERENT NEURODEGENERATIVE DISEASES

A cross-brain-regions study of *ANK1* DNA methylation in different neurodegenerative diseases.

Adam R Smith[1], Rebecca G Smith[1], Joe Burrage[1], Claire Troakes[2], Safa Al-Sarraj[2], Rajesh N Kalaria[3], Carolyn Sloan[4], Andrew C Robinson[5], Jonathan Mill[1], Katie Lunnon[1,*]

[1] University of Exeter Medical School, University of Bristol, Exeter, UK.

[2] Institute of Psychiatry, Psychology and Neuroscience (IoPPN), King's College London, London, UK.

[3] Institute of Neuroscience, Newcastle University, Newcastle, UK.

[4] Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK.

[5] Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK.

* Corresponding author: Katie Lunnon, University of Exeter Medical School, RILD Building Level 4, Royal Devon and Exeter Hospital, Barrack Rd, Exeter. EX2 5DW. UK. E-mail: k.lunnon@exeter.ac.uk

ABSTRACT

Recent epigenome-wide association studies in Alzheimer's disease have highlighted consistent robust neuropathology-associated DNA hypermethylation of the Ankyrin 1 (*ANK1*) gene in the cortex. The extent to which altered *ANK1* DNA methylation is also associated with other neurodegenerative diseases is not currently known. In the current study, we used bisulfite pyrosequencing to quantify DNA methylation across eight CpG sites within a 118bp region of the *ANK1* gene across multiple brain regions in Alzheimer's disease, Vascular dementia, Dementia with Lewy bodies, Huntington's disease and Parkinson's disease. We demonstrate disease-associated *ANK1* hypermethylation in the entorhinal cortex in Alzheimer's disease, Huntington's disease and Parkinson's disease, whilst in donors with Vascular dementia and Dementia with Lewy bodies we observed elevated *ANK1* DNA methylation only in individuals with co-existing Alzheimer's disease pathology. We did not observe any disease-associated differential *ANK1* DNA methylation in the striatum in Huntington's disease, or the substantia nigra in Parkinson's disease. Our data suggests that *ANK1* is characterized by region and disease-specific differential DNA methylation in multiple neurodegenerative diseases.

Keywords:

Alzheimer's disease (AD); Ankyrin 1 (ANK1); Brain; Dementia with Lewy Bodies (DLB); DNA methylation (5-methylcytosine – 5mC); Epigenetics; Huntington's disease (HD); Parkinson's disease (PD); Vascular dementia (VaD)

1. INTRODUCTION

Dementia encompasses a group of chronic neurodegenerative diseases that affected an estimated 46.8 million people worldwide in 2015 (Wimo, et al., 2017), of which Alzheimer's disease (AD) accounts for ~60% of cases. The etiology of AD has been hypothesized to involve epigenetic mechanisms (Lunnon and Mill, 2013). In 2014 two epigenome-wide association studies (EWAS) of AD identified significant hypermethylation of CpG sites in the Ankyrin 1 (*ANK1*) gene associated with neuropathology in AD cortex (De Jager, et al., 2014,Lunnon, et al., 2014), which has been replicated in multiple independent study cohorts (Smith, 2017). Subsequently, a genome wide association study (GWAS) of a Han Chinese population identified a single-nucleotide polymorphism (SNP) in *ANK1* associated with an increased susceptibility for developing AD (Chi, et al., 2015). *ANK1* links integral membrane proteins to the underlying spectrin-actin cytoskeleton and plays a key role in cell motility, activation, proliferation, contact, and maintenance of specialized membrane domains (Yang, et al., 2011). There is now increasing interest in understanding the role of epigenetic changes in *ANK1* in the development and progression of AD. One important question to be addressed is whether *ANK1* hypermethylation is specific to AD, or observed in other neurodegenerative disorders. Although AD accounts for ~60% of dementia cases, many other dementias share common symptoms and/or pathological hallmarks with AD.

This study aimed to quantify DNA methylation levels across a 118bp region of *ANK1*, previously associated with AD, in a number of different neurodegenerative diseases. Using bisulfite pyrosequencing we assessed *ANK1* DNA methylation in brain samples from donors with AD, Dementia with Lewy bodies (DLB), Vascular dementia (VaD), Huntington's disease (HD), Parkinson's disease (PD) and non-demented

elderly controls, across a number of different brain regions that are characterized by disease-specific pathology.

2. MATERIALS AND METHODS

2.1. Subjects and Samples

Post-mortem brain tissue was obtained from six different UK brain banks (the South West Dementia Brain Bank (SWDBB), the London Neurodegenerative Disease Brain Bank (LNDBB) the Manchester Brain Bank, the Oxford Brain Bank, the Cambridge Brain Bank and the Newcastle Brain Bank). In total, tissue was obtained from 60 AD (Braak V-VI), 119 DLB, 27 VaD, 22 HD, 36 PD and 105 elderly non-demented control subjects (Braak 0-II). A subset of DLB (N = 39) and VaD (N = 5) cases also had co-existing AD pathology. For each disease we analyzed the entorhinal cortex (EC), superior temporal gyrus (STG) and cerebellum (CER). For HD cases we also analyzed the striatum (STR) as this is primarily affected in disease (Reiner, et al., 2011), whilst for PD cases we analyzed the STR and substantia nigra (SN) as these are regions of pathology in this disease (Fearnley and Lees, 1991). For control samples we analyzed all five brain regions. For a small number of donors tissue was not available from all brain regions. Genomic DNA was isolated from ~100mg of each dissected brain region using a standard phenol-chloroform extraction method, and tested for degradation and purity prior to analysis as previously described (Smith, et al., 2016). Demographic information for samples can be found in **Supplementary Table 1**.

2.2. ANK1 bisulfite pyrosequencing

Bisulfite pyrosequencing was used to quantify DNA methylation across eight individual CpG sites in the *ANK1* gene, spanning from 41519302 to 41519420 within

chromosome 8 (hg19). Bisulfite conversion was performed using the Bisulfite-Gold kit (Zymo research, USA). A single amplicon (246bp) was generated using primers designed using the PyroMark Assay Design software 2.0 (Qiagen, UK) as previously described (Lunnon, et al., 2014). Pyrosequencing was performed using two sequencing primers to maximize coverage across eight CpG sites. DNA methylation was quantified using the Pyromark Q24 system (Qiagen, UK) following the manufacturer's standard instructions and the Pyro Q24 CpG 2.0.6 software.

2.3. Data Analysis

All computations and statistical analyses were performed using R 3.3.2 (R Development Core Team, 2012). A linear regression analysis was performed, controlling for the effects of age, gender and batch effects, comparing control samples with samples affected by each neurodegenerative disease. For the VaD and DLB samples we also performed a second analysis to investigate whether co-existing AD pathology influenced the results by comparing individuals with and without co-existing AD pathology to control samples, again accounting for the effects of age, gender and batch. Our analyses examined DNA methylation differences at (a) individual CpG sites and (b) averaged across the amplicon. We used a paired two-tailed t-test to compare adjusted DNA methylation differences in disease across brain regions.

3. RESULTS

3.1. AD-associated ANK1 DNA hypermethylation is seen across all tissues analyzed

First, we sought to replicate previous findings of *ANK1* DNA hypermethylation in AD. Across the 118bp region, we observed significantly increased levels in AD cases

compared to controls in all eight *ANK1* CpG sites in the EC (**Figure 1A**) and seven *ANK1* CpG sites in the STG (**Figure 1B**). Both the EC and STG exhibit a high degree of AD pathology, even in the earliest stages of disease, with the EC being the starting point of AD pathology in the cortex (Braak and Braak, 1991). Conversely, the CER remains free of AD pathology until the very last stages of the disease, although even then this is limited to amyloid beta (Aβ) plaques with an absence of neurofibrillary tangles (NFTs) of hyperphosphorylated tau (Braak, et al., 1989). Of note, we observed significant *ANK1* DNA hypermethylation at five *ANK1* CpG sites in the CER (**Figure 1C**). This is the first time *ANK1* DNA methylation changes have been reported in the CER. Interestingly, two of the loci that did not display AD-associated *ANK1* hypermethylation in the CER were chr8:41519308 and chr8:41519399, the two sites included on the Illumina 450K array used in previous EWAS analyses of AD which did not identify *ANK1* hypermethylation in AD in the CER (Lunnon, et al., 2014). Average DNA methylation across the amplicon region was significantly elevated in AD in the EC (P = 1.29 x 10^{-07}), STG (P = 2.39 x 10^{-03}) and CER (P = 7.81 x 10^{-03}) (**Figure 1D**). *ANK1* DNA methylation differences between cases and controls at both individual sites and across the amplicon were lower in the CER compared to other tissues tested (**Supplementary Table 2**), with a significantly greater DNA methylation difference between cases and controls in the EC (amplicon average corrected DNA methylation difference (Δ) = 4.53%) compared to both the STG (amplicon average Δ = 2.84%; P = 7.98 x 10^{-4}) and the CER (amplicon average Δ = 1.17%, P = 2.55 x 10^{-4}). Interestingly, this pattern of change matches the spread of AD pathology throughout the brain.

3.2. ANK1 DNA hypermethylation in the EC is only observed in DLB cases with co-existing AD pathology

DLB is the third most common cause of dementia with the age of onset ranging from 50 to 83 years (McKeith, 2002). The pathology of DLB shares similarities to AD, with

the presence of immune regulation and microglial activation being consistent between diseases (Mackenzie, 2000). However, the presence of Lewy bodies within the brain makes DLB considerably more comparable to PD (McKeith, 2002). In fact, PD dementia is thought to be biologically identical to DLB, only differing in the order in which the motor or cognitive symptoms occur (Dodel, et al., 2008). Interestingly we observed significant hypermethylation of *ANK1* in DLB cases compared to controls in the EC (**Figure 2A**) at four of the eight *ANK1* CpG sites **(Supplementary Table 3)**. We saw no difference between DLB and control samples in either the STG (**Supplementary Figure 1A**) or the CER (**Supplementary Figure 1B**) at any of the eight *ANK1* CpG sites. Across the *ANK1* amplicon we observed significant DLB-associated hypermethylation in the EC ($P = 0.0244$), but not in the STG or CER ($P > 0.05$) (**Figure 2B**). It is widely reported that DLB and AD frequently co-occur (Rosenberg, et al., 2001); we were therefore interested to investigate whether we still observed DLB-associated DNA hypermethylation in the EC when we controlled for co-existing AD pathology. We found no significant changes in *ANK1* DNA methylation in individuals with "pure" DLB compared to controls in the EC (**Figure 2C**), STG (**Supplementary Figure 1C**) and CER (**Supplementary Figure 1D**). However, we did observe significant hypermethylation in DLB cases with co-existing AD pathology compared to controls at seven of the eight *ANK1* CpG sites in the EC **(Figure 2C)** and two sites in the STG **(Supplementary Figure 1C),** with no difference in the CER (**Supplementary Figure 1D**). When we looked across the whole 118bp region, we saw increased *ANK1* DNA methylation in the EC in individuals with co-existing AD pathology ($P = 1.45 \times 10^{-03}$) **(Figure 2D**), suggesting that the *ANK1* hypermethylation seen in some individuals with DLB is primarily driven by AD pathology.

3.3. ANK1 hypermethylation is seen in the EC only in VaD individuals with co-existing AD pathology

Characterized by the loss of neurological function due to ischemic events, the risk of developing VaD is closely linked to vascular health (Román, et al., 1993). We observed increased DNA methylation in individuals with VaD at none of the ANK1 CpG sites in the EC (**Figure 3A**), only one site in the STG (**Supplementary Figure 2A**) and no sites in the CER (**Supplementary Figure 2B**) (**Supplementary Table 4**), with no difference across the amplicon in any of the brain regions tested ($P > 0.05$) (**Figure 3B**). Because VaD also often co-occurs with AD we next examined whether stratifying cases by the presence of AD pathology altered these findings. Interestingly, we saw disease-associated hypermethylation in the EC at five of the eight ANK1 CpG sites only in individuals with co-existing AD pathology (**Figure 3C**), whilst we saw disease-associated hypomethylation at one site in the STG (**Supplementary Figure 2C)** in individuals with "pure" VaD, and no disease-associated changes in the CER (**Supplementary Figure 2D**). When we looked across the 118bp region we only saw significant *ANK1* hypermethylation in individuals with VaD and co-existing AD pathology compared to controls in the EC ($P = 0.0163$) **(Figure 3D)**. It is worth noting that our cohort only had a small number of VaD cases with co-existing AD pathology (N=5).

3.4. ANK1 DNA hypermethylation in the EC is seen in both HD and PD.

HD is characterized by a trinucleotide repeat in the huntingtin gene (*HTT*). The abundance of the repeat is proportional to the level of protein misfolding and downstream cytosolic accumulation, leading to neuronal cell death and the symptoms of HD (Walker, 2007). *ANK1* DNA hypermethylation was seen at four of the eight CpG sites in the EC in HD **(Figure 4A**). However, no differential DNA methylation was seen in the other brain regions tested (**Supplementary Table 5**), including the STG (**Supplementary Figure 3A**), the CER (**Supplementary Figure 3B**) and the STR (**Supplementary Figure 3C**), a region that forms part of the basal ganglia, known to be the first brain region to be adversely affected by HD pathology

(Walker, 2007). Averaging across the region again highlighted significant hypermethylation in the EC (P = 6.68 x 10^{-3}), with no significant change in any other tissue **(Figure 4B)**.

A similar pattern of *ANK1* hypermethylation was observed in PD; two of the eight CpG sites were characterized by significant hypermethylation in the EC (**Figure 4C**), with no differences in DNA methylation in any of the other brain regions tested (**Supplementary Table 6**). This included the STG (**Supplementary Figure 4A**), the CER (**Supplementary Figure 4B**), the STR (**Supplementary Figure 4C**) and the SN (**Supplementary Figure 4D**), with the SN representing the brain region that has the highest levels of pathology in PD (Fearnley and Lees, 1991). Across the 118bp amplicon we saw no change in DNA methylation in any of the five brain regions (**Figure 4D**).

4. DISCUSSION

This is the first study to assess brain *ANK1* DNA methylation changes across multiple neurodegenerative diseases. We identified significant DNA methylation changes in the EC in multiple diseases, including AD, HD and PD, with significant DNA hypermethylation across the amplicon in AD and HD. Interestingly, we also observed significant hypermethylation of *ANK1* in the EC in both DLB and VaD at several individual CpG sites and across the amplicon, but only in donors with co-existing AD pathology. This suggests that *ANK1* DNA hypermethylation in the EC is specific to some neurodegenerative diseases (AD, HD and to some extent PD), and not observed in other forms of neuropathology (VaD and DLB). We have previously reported that *ANK1* is not hypermethylated in the CER in AD at two sites interrogated by the Illumina 450K array (chr8:41519308 and chr8:41519399). In the current study

we again demonstrate that these two loci are not significantly differentially methylated in AD; however, we do highlight AD-associated DNA hypermethylation at five adjacent CpG sites and averaged across the 118bp amplicon in the CER. We did not see any *ANK1* DNA methylation changes in the CER in any of the other neurodegenerative diseases, including those with co-existing AD pathology. Reflecting our previous findings, we found that DNA methylation differences in AD are greatest in the EC, an area with high levels of neuropathology and lowest in the CER, the region with the least neuropathology. Interestingly, although we observed disease-associated *ANK1* hypermethylation in the EC at six individual sites and across the region in HD and at two individual sites in PD, we did not see *ANK1* DNA methylation changes in these diseases in their regions of primary neuropathology, namely the STR and SN, respectively. This suggests that *ANK1* hypermethylation in neurodegenerative disease is not necessarily specific to regions of primary neuropathology, but may instead be specific to particular cell types affected in the diseases, such as those in the EC, which are not present in the STR and SN. *ANK1* encodes for numerous isoforms with their own tissue-specific enhancers. Although the precise function of most *ANK1* isoforms is not known, different isoforms have been identified in the brain, blood and muscle (Gallagher, et al., 1997). Mastroeni *et al*., recently showed a four-fold increase in *ANK1* mRNA expression in microglia from AD brain tissue, but not in neurons or astrocytes from the same individuals, suggesting an immune based function for ANK1 in the human brain (Mastroeni, et al., 2017). One potential caveat of our study is that we have analyzed “bulk” tissue, and we cannot determine which cell type(s) are driving the DNA hypermethylation seen in *ANK1* in disease.

Although the focus of our study was on investigating DNA methylation changes in disease, bisulfite pyrosequencing actually generates a summative measurement of both DNA methylation and DNA hydroxymethylation. DNA methylation is generally

associated with gene silencing, particularly when residing in the promoters of constitutively expressed housekeeping genes (Jones, 2012), whilst DNA hydroxymethylation has been shown to be enriched in gene bodies (Lunnon, et al., 2016) and to be found at (relatively) high levels in the brain (Khare, et al., 2012,Szulwach, et al., 2011). We have recently shown that *ANK1* DNA hypermethylation across the 118bp amplicon in AD is actually underestimated when using bisulfite data, as it is confounded by significant DNA hypohydroxymethylation at some loci in the amplicon (Smith, et al., Under Review). Another caveat to our study is that we have only analyzed DNA methylation across eight CpG sites in a 118bp region of the *ANK1* gene and thus future studies should aim to further quantify changes in DNA methylation across the entire 244kb gene.

5. CONCLUSIONS

Our study has demonstrated disease-associated *ANK1* hypermethylation in the EC at specific CpG sites in AD, HD and PD and across the region in AD and HD. In donors with DLB and VaD we only observed increased *ANK1* DNA methylation in the EC in individuals with co-existing AD pathology. The CER showed disease-associated hypermethylation at specific CpG loci and across the region in AD, but not in any of the other neurodegenerative diseases tested. We saw no disease-associated differential *ANK1* DNA methylation in the STR in HD or PD, or the SN in PD. This suggests that *ANK1* is characterized by brain region and disease-specific differential DNA methylation in different neurodegenerative diseases. At present we are unable to determine whether these changes represent a cause or a consequence of the disease process. Further functional studies should therefore be performed to determine the potential disease causality of this modification.

ACKNOWLEDGEMENTS

This work was funded by an Alzheimer's Association US New Investigator Research Grant (NIRG-14-320878) to KL, a grant from BRACE (Bristol Research into Alzheimer's and Care of the Elderly) to KL, a major project grant from the Alzheimer's Society UK (AS-PG-14-038) to KL and an Equipment Grant from Alzheimer's Research UK (ART-EG2010A-2) to JM. We thank all the donors and families who have made this research possible. Brain tissue was received from six of the UK Brain Banks. Brain tissue collection by the Newcastle brain bank, Oxford Brain Bank, The London Neurodegenerative Diseases Brain Bank (LNDBB) and the South West Dementia Brain Bank (SWDBB) is supported by the UK Medical Research Council (G0400074). The Newcastle Brain Bank, Oxford Brain Bank, LNDBB, SWDBB and Manchester Brain Bank are all supported by the Brains for Dementia Research (BDR) program, jointly funded by Alzheimer's Research UK and Alzheimer's Society. In addition, the Newcastle Brain Bank is also supported by the NIHR Newcastle Biomedical Research Centre and Unit award to the Newcastle upon Tyne NHS Foundation Trust and Newcastle University. The Oxford Brain Bank is also supported by Autistica UK and the NIHR Oxford Biomedical Research Centre. The SWDBB is also supported by BRACE. The Cambridge Brain Bank is supported by the NIHR and the Cambridge Biomedical Research Centre.

REFERENCES

Braak, H., Braak, E. 1991. Neuropathological stageing of Alzheimer-related changes. Acta neuropathologica 82(4), 239-59.

Braak, H., Braak, E., Bohl, J., Lang, W. 1989. Alzheimer's disease: amyloid plaques in the cerebellum. Journal of the Neurological Sciences 93(2), 277-87. doi:https://doi.org/10.1016/0022-510X(89)90197-4.

Chi, S., Song, J.H., Tan, M.S., Zhang, W., Wang, Z.X., Jiang, T., Tan, L., Yu, J.T. 2015. Association of Single-Nucleotide Polymorphism in ANK1

with Late-Onset Alzheimer's Disease in Han Chinese. Mol Neurobiol. doi:10.1007/s12035-015-9547-x.

De Jager, P.L., Srivastava, G., Lunnon, K., Burgess, J., Schalkwyk, L.C., Yu, L., Eaton, M.L., Keenan, B.T., Ernst, J., McCabe, C., Tang, A., Raj, T., Replogle, J., Brodeur, W., Gabriel, S., Chai, H.S., Younkin, C., Younkin, S.G., Zou, F., Szyf, M., Epstein, C.B., Schneider, J.A., Bernstein, B.E., Meissner, A., Ertekin-Taner, N., Chibnik, L.B., Kellis, M., Mill, J., Bennett, D.A. 2014. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. Nat Neurosci 17(9), 1156-63. doi:10.1038/nn.3786 http://www.nature.com/neuro/journal/v17/n9/abs/nn.3786.html#supplementary-information.

Dodel, R., Csoti, I., Ebersbach, G., Fuchs, G., Hahne, M., Kuhn, W., Oechsner, M., Jost, W., Reichmann, H., Schulz, J.B. 2008. Lewy body dementia and Parkinson's disease with dementia. Journal of neurology 255 Suppl 5, 39-47. doi:10.1007/s00415-008-5007-0.

Fearnley, J.M., Lees, A.J. 1991. Aging and Parkinson's disease: substantia nigra regional selectivity. Brain : a journal of neurology 114(5), 2283-301. doi:10.1093/brain/114.5.2283.

Gallagher, P.G., Tse, W.T., Scarpa, A.L., Lux, S.E., Forget, B.G. 1997. Structure and Organization of the Human Ankyrin-1 Gene: BASIS FOR COMPLEXITY OF PRE-mRNA PROCESSING. Journal of Biological Chemistry 272(31), 19220-8. doi:10.1074/jbc.272.31.19220.

Jones, P.A. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature reviews Genetics 13. doi:10.1038/nrg3230.

Khare, T., Pai, S., Koncevicius, K., Pal, M., Kriukiene, E., Liutkeviciute, Z., Irimia, M., Jia, P., Ptak, C., Xia, M., Tice, R., Tochigi, M., Morera, S., Nazarians, A., Belsham, D., Wong, A.H., Blencowe, B.J., Wang, S.C., Kapranov, P., Kustra, R., Labrie, V., Klimasauskas, S., Petronis, A. 2012. 5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary. Nature structural & molecular biology 19(10), 1037-43. doi:10.1038/nsmb.2372.

Lunnon, K., Hannon, E., Smith, R.G., Dempster, E., Wong, C., Burrage, J., Troakes, C., Al-Sarraj, S., Kepa, A., Schalkwyk, L., Mill, J. 2016. Variation in 5-hydroxymethylcytosine across human cortex and cerebellum. Genome biology 17, 27. doi:10.1186/s13059-016-0871-x.

Lunnon, K., Mill, J. 2013. Epigenetic studies in Alzheimer's disease: current findings, caveats, and considerations for future studies. Am J Med Genet B Neuropsychiatr Genet 162B(8), 789-99. doi:10.1002/ajmg.b.32201.

Lunnon, K., Smith, R., Hannon, E., De Jager, P.L., Srivastava, G., Volta, M., Troakes, C., Al-Sarraj, S., Burrage, J., Macdonald, R., Condliffe, D., Harries, L.W., Katsel, P., Haroutunian, V., Kaminsky, Z., Joachim, C., Powell, J., Lovestone, S., Bennett, D.A., Schalkwyk, L.C., Mill, J. 2014. Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. Nature neuroscience 17(9), 1164-70. doi:10.1038/nn.3782.

Mackenzie, I.R.A. 2000. Activated microglia in dementia with Lewy bodies. Neurology 55(1), 132-4. doi:10.1212/wnl.55.1.132.

Mastroeni, D., Sekar, S., Nolz, J., Delvaux, E., Lunnon, K., Mill, J., Liang, W.S., Coleman, P.D. 2017. ANK1 is up-regulated in laser captured microglia in Alzheimer's brain; the importance of addressing cellular heterogeneity. PloS one 12(7), e0177814. doi:10.1371/journal.pone.0177814.

McKeith, I.G. 2002. Dementia with Lewy bodies. The British Journal of Psychiatry 180(2), 144-7. doi:10.1192/bjp.180.2.144.

R Development Core Team. 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria 2012.

Reiner, A., Dragatsis, I., Dietrich, P. 2011. GENETICS AND NEUROPATHOLOGY OF HUNTINGTON'S DISEASE. International review of neurobiology 98, 325-72. doi:10.1016/B978-0-12-381328-2.00014-6.

Román, G.C., Tatemichi, T.K., Erkinjuntti, T., Cummings, J.L., Masdeu, J.C., Garcia, J.H., Amaducci, L., Orgogozo, J.-M., Brun, A., Hofman, A., Moody, D.M., O'Brien, M.D., Yamaguchi, T., Grafman, J., Drayer, B.P., Bennett, D.A., Fisher, M., Ogata, J., Kokmen, E., Bermejo, F., Wolf, P.A., Gorelick, P.B., Bick, K.L., Pajeau, A.K., Bell, M.A., DeCarli, C., Culebras, A., Korczyn, A.D., Bogousslavsky, J., Hartmann, A., Scheinberg, P. 1993. Vascular dementia. Diagnostic criteria for research studies: Report of the NINDS-AIREN International Workshop* 43(2), 250-. doi:10.1212/wnl.43.2.250.

Rosenberg, C.K., Cummings, T.J., Saunders, A.M., Widico, C., McIntyre, L.M., Hulette, C.M. 2001. Dementia with Lewy bodies and Alzheimer's disease. Acta Neuropathol 102(6), 621-6. doi:10.1007/s004010100415.

Smith, A.R., Smith, R.G., Condliffe, D., Hannon, E., Schalkwyk, L., Mill, J., Lunnon, K. 2016. Increased DNA methylation near TREM2 is consistently seen in the superior temporal gyrus in Alzheimer's disease brain. Neurobiology of aging 47, 35-40. doi:10.1016/j.neurobiolaging.2016.07.008.

Smith, A.R., Smith, R.G., Hannon, E., Roubroeks, J.A.Y., Burrage, J., Troakes, C., Al-Sarraj, S., Mill, J., van den Hove, D.L., Lunnon, K. Under Review. Parallel profiling of DNA methylation and hydroxymethylation highlights neuropathology-associated epigenetic variation within ANK1 and thirteen novel loci in Alzheimer's disease entorhinal cortex. Genome biology.

Smith, R.G.L., K. 2017. DNA Modifiactions and Alzheimer's Disease. in: Delgado-Morales, R. (Ed.). Neuroepigenomics in Aging and Disease Springer International Publishing, pp 303-19.

Szulwach, K.E., Li, X., Li, Y., Song, C.X., Wu, H., Dai, Q., Irier, H., Upadhyay, A.K., Gearing, M., Levey, A.I., Vasanthakumar, A., Godley, L.A., Chang, Q., Cheng, X., He, C., Jin, P. 2011. 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. Nature neuroscience 14(12), 1607-16. doi:10.1038/nn.2959.

Walker, F.O. 2007. Huntington's disease. The Lancet 369(9557), 218-28. doi:https://doi.org/10.1016/S0140-6736(07)60111-1.

Wimo, A., Guerchet, M., Ali, G.C., Wu, Y.T., Prina, A.M., Winblad, B., Jonsson, L., Liu, Z., Prince, M. 2017. The worldwide costs of dementia 2015 and comparisons with 2010. Alzheimer's & dementia : the journal

of the Alzheimer's Association 13(1), 1-7. doi:10.1016/j.jalz.2016.07.150.

Yang, M., Ge, W., Chowdhury, R., Claridge, T.D., Kramer, H.B., Schmierer, B., McDonough, M.A., Gong, L., Kessler, B.M., Ratcliffe, P.J., Coleman, M.L., Schofield, C.J. 2011. Asparagine and aspartate hydroxylation of the cytoskeletal ankyrin family is catalyzed by factor-inhibiting hypoxia-inducible factor. The Journal of biological chemistry 286(9), 7648-60. doi:10.1074/jbc.M110.193540.

Figure 1: *ANK1* is hypermethylated in the EC, STG and CER in AD brain. Using bisulfite pyrosequencing we assayed a 118bp region of the *ANK1* gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in (**A**) EC, (**B**) STG and (**C**) CER tissue. in AD samples (Braak V-VI) compared to control samples (Braak 0-II). We demonstrated significant neuropathology-associated hypermethylation at all assayed CpG sites in the EC, significant hypermethylation at seven of the eight CpG sites in the STG and five of the eight sites in the CER. (**D**) When data was summed across the 118bp amplicon region we observed a significant increase in DNA methylation across all brain regions (EC: $P = 1.29 \times 10^{-7}$, STG: $P = 2.39 \times 10^{-3}$ and CER: $P = 7.81 \times 10^{-3}$). Data is represented as mean (±SEM) Key: $*P < 0.05$, $**P < 0.01$, $***P < 0.005$.

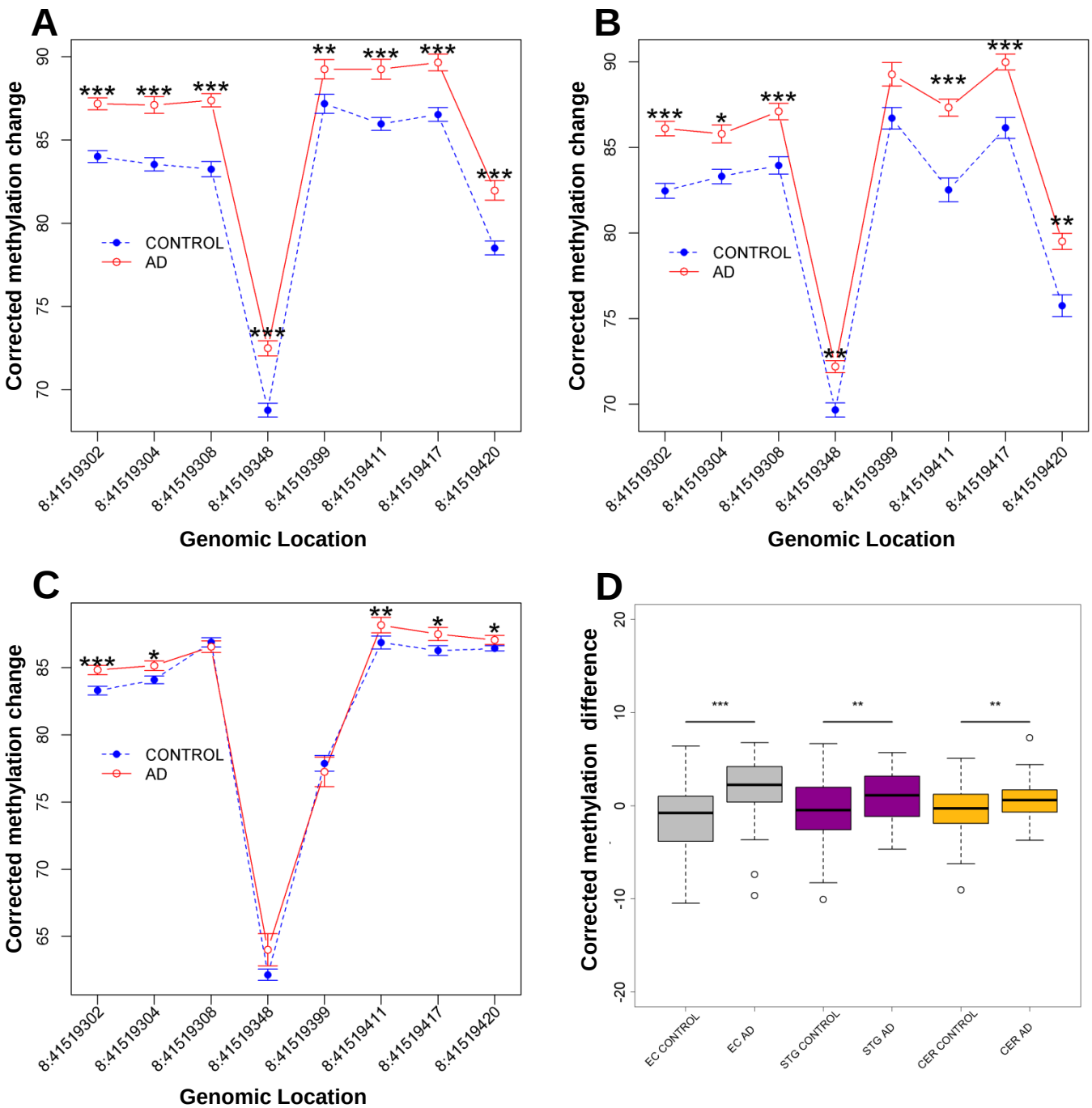


Figure 2: *ANK*1 hypermethylation is observed in the EC in individuals with DLB and co-existing AD pathology. Using bisulfite pyrosequencing we assayed a 118bp region of the *ANK1* gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in all DLB samples compared to control samples. (**A**) We demonstrated significant neuropathology-associated hypermethylation at four CpG sites in the EC. (**B**) When data was summed across the 118bp amplicon region we observed a significant increase in DNA methylation in the EC (*P* = 0.024). Some individuals with DLB also had co-existing AD pathology; (**C**) when we compared DNA methylation levels in ANK1 in individuals with DLB and co-existing AD pathology, or individuals with "pure" DLB only to controls, we found significant hypermethylation at seven sites in the EC in individuals with co-existing AD pathology. (**D**) When we averaged methylation across the region we saw significant hypermethylation in the EC in individuals with co-existing AD pathology *(P* = 0.001). Data is represented as mean (±SEM) Key: **P* < 0.05, ***P* < 0.01, ****P* < 0.005.

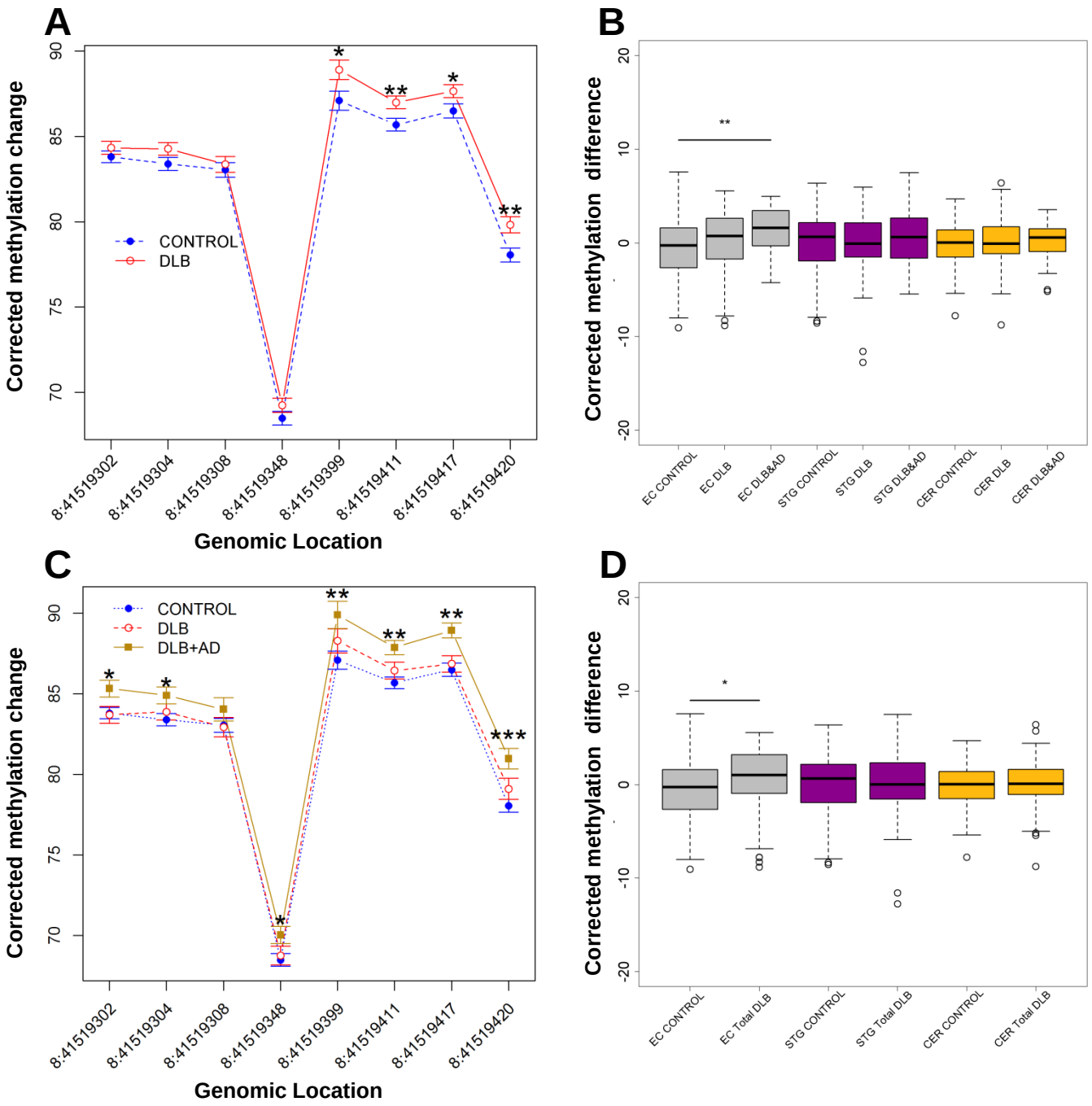


Figure 3: *ANK1* hypermethylation is observed in the EC in individuals with VaD and co-existing AD pathology. Using bisulfite pyrosequencing we assayed a 118bp region of the *ANK1* gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in all VaD samples compared to control samples. (**A**) We found no disease-associated differential methylation at any individual sites in the EC, (**B**) nor any difference when averaged across the amplicon (**B**). Some individuals with VaD also had co-existing AD pathology; (**C**) when we compared DNA methylation levels in *ANK1* in individuals with VaD and co-existing AD pathology, or individuals with "pure" VaD only to controls, we found significant hypermethylation at five sites in the EC in individuals with co-existing AD pathology. (**D**) When we averaged methylation across the region we observed significant hypermethylation in the EC in individuals with a co-diagnosis of AD *(P* = 0.016). Data is represented as mean (±SEM) Key: **P* < 0.05, ***P* < 0.01, ****P* < 0.005.

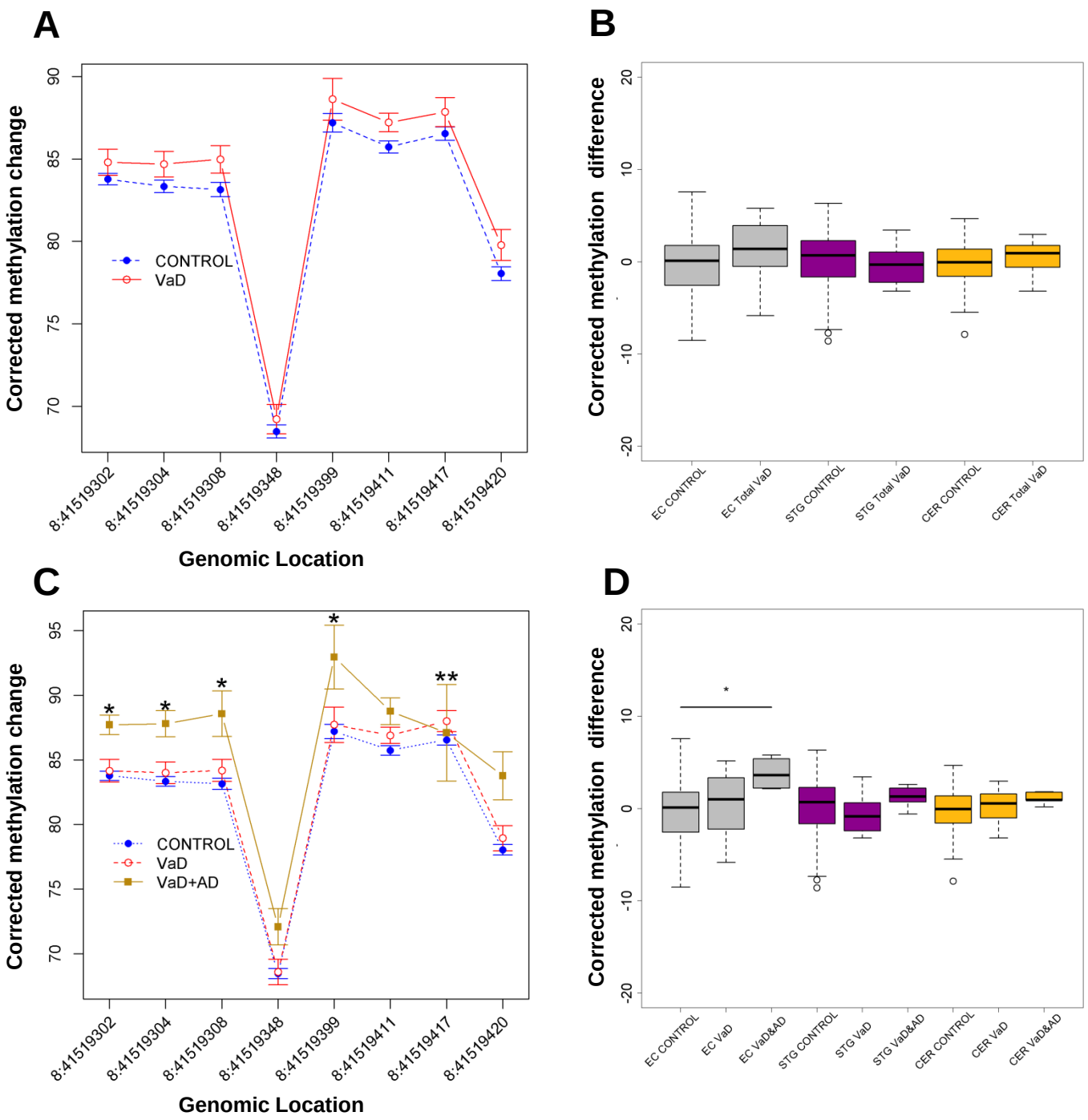
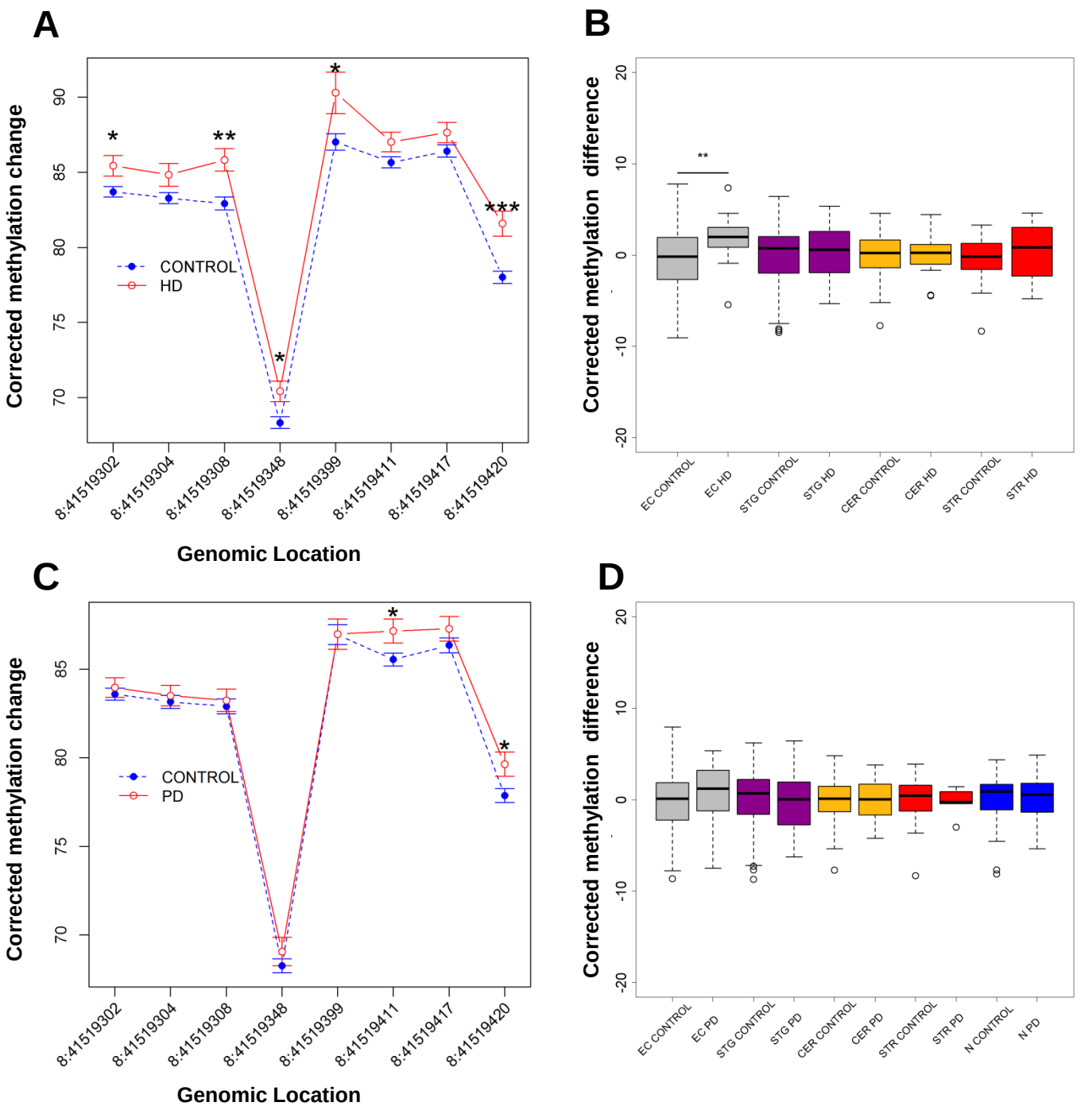


Figure 4: *ANK1* DNA methylation patterns in HD and PD. Using bisulfite pyrosequencing we assayed a 118bp region of the *ANK1* gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in HD and PD samples compared to control samples. In HD donors in the EC we observed neuropathology-associated hypermethylation at five sites (**A**) and significant hypermethylation when we averaged methylation across the region *(P* = 0.007) (**B**). In PD donors in the EC we found neuropathology-associated hypermethylation at two sites (**C**), but no significant difference when we averaged across the region (**D**). Data is represented as mean (±SEM) Key: **P* < 0.05, ***P* < 0.01, ****P* < 0.005.



Supplementary Table 1: Sample and demographic information for samples used in the study. Shown for each disease cohort are the number of total donors used (N), then for each individual brain region, the number of samples available (N), gender distribution and mean age (± standard deviation (SD)). Abbreviations: Alzheimer's disease (AD), Cerebellum (CER), Dementia with Lewy bodies (DLB), Entorhinal cortex (EC), Huntington's disease (HD), Parkinson's disease (PD), Substantia nigra (SN), Superior temporal gyrus (STG), Striatum (STR), Vascular dementia (VaD).

		Brain Region				
		EC	**STG**	**CER**	**STR**	**SN**
AD (N = 60)	**N**	58	58	58	-	-
	Gender (M/F)	27/31	26/32	26/32	-	-
	Mean age (±SD)	77.7 (9.2)	77.6 (9.2)	78.4 (8.8)	-	-
All DLB (N = 119)	**N**	91	107	112	-	-
	Gender (M/F)	72/19	84/23	88/24	-	-
	Mean age (±SD)	78.8 (7.3)	78.3 (7.2)	78.1 (7.2)	-	-
DLB only (N = 80)	**N**	56	69	76	-	-
	Gender (M/F)	47/9	56/13	61/15	-	-
	Mean age (±SD)	78.1 (7.1)	77.6 (6.9)	77.4 (7.0)	-	-
DLB + AD (N = 39)	**N**	35	38	36	-	-
	Gender (M/F)	25/10	28/10	27/9	-	-
	Mean age (±SD)	80.0 (7.5)	79.5 (7.6)	79.5 (7.7)	-	-
All VaD (N = 27)	**N**	23	27	26	-	-
	Gender (M/F)	9/14	11/16	11/15	-	-
	Mean age (±SD)	83.9 (8.4)	83.6 (7.9)	83.5 (8.1)	-	-
VaD only (N = 22)	**N**	19	22	21	-	-
	Gender (M/F)	8/11	10/12	10/11	-	-
	Mean age (±SD)	83.3 (7.8)	83.0 (7.4)	82.8 (7.5)	-	-
VaD + AD (N = 5)	**N**	4	5	5	-	-
	Gender (M/F)	1/3	1/4	1/4	-	-
	Mean age (±SD)	86.8 (12.1)	86.2 (10.6)	86.2 (10.6)	-	-
HD (N = 22)	**N**	19	20	20	19	-
	Gender (M/F)	13/6	13/7	14/6	13/6	-
	Mean age (±SD)	71.1 (5.0)	70.4 (7.1)	70.0 (7.2)	68.8 (12.8)	-
PD (N = 36)	**N**	33	35	33	5	29
	Gender (M/F)	19/14	21/14	21/12	4/1	19/10
	Mean age (±SD)	78.5 (7.4)	78.2 (7.5)	78.3 (7.5)	83.6 (3.2)	78.9 (7.6)
Control (N = 105)	**N**	99	74	98	21	31
	Gender (M/F)	53/46	40/34	48/50	9/12	11/20
	Mean age (±SD)	81.4 (9.6)	81.6 (10.6)	80.7 (10.7)	76.5 (11.7)	77.3 (11.2)

Supplementary Table 2: *ANK1* CpG sites are hypermethylated in AD across all brain regions tested. Shown for each tissue is chromosomal location (genome build hg19) of CpG sites tested, with corrected difference (Δ) in DNA methylation levels and corresponding *P* value between controls (Braak 0-II) and AD cases (Braak V-VI) after adjusting for the covariates of age, sex and batch effects. A positive Δ reflects hypermethylation in AD. All *P* values < 0.05 (and their corresponding Δ) are shown in bold.

Genomic Location	AD					
	EC		STG		CER	
	Δ	***P***	Δ	***P***	Δ	***P***
chr8:41519302	**4.31**	**1.76E-10**	**2.83**	**1.49E-04**	**2.09**	**2.98E-04**
chr8:41519304	**4.85**	**5.76E-10**	**2.01**	**0.017**	**1.42**	**0.008**
chr8:41519308	**5.51**	**3.16E-11**	**3.22**	**4.43E-04**	-0.23	0.729
chr8:41519348	**4.99**	**2.79E-10**	**2.33**	**0.002**	0.17	0.838
chr8:41519399	**2.87**	**0.005**	1.68	0.150	0.66	0.585
chr8:41519411	**4.58**	**2.27E-08**	**4.02**	**2.22E-04**	**2.83**	**8.43E-04**
chr8:41519417	**4.30**	**1.50E-07**	**3.53**	**4.55E-04**	**1.54**	**0.028**
chr8:41519420	**4.80**	**2.53E-08**	**3.08**	**0.003**	**0.86**	**0.038**
Amplicon Average	**3.14**	**1.29E-07**	**1.68**	**0.002**	**0.98**	**0.008**

Supplementary Table 3: *ANK1* CpG sites are hypermethylated in the EC in patients with DLB and co-existing AD pathology. Shown for each tissue is chromosomal location (genome build hg19) of CpG sites tested, with corrected difference (Δ) in DNA methylation levels and corresponding *P* value between controls and DLB cases after adjusting for the covariates of age, sex and batch effects. Analyses are stratified into "All DLB", "DLB Only" and "DLB + AD" groups. A positive Δ reflects hypermethylation in disease. All *P* values < 0.05 (and their corresponding Δ) are shown in bold.

	All DLB						DLB Only						DLB + AD					
	EC		STG		CER		EC		STG		CER		EC		STG		CER	
Genomic Location	Δ	*P*	Δ	*P*	Δ	*P*	Δ	*P*	Δ	*P*	Δ	*P*	Δ	*P*	Δ	*P*	Δ	*P*
chr8:41519302	0.58	0.286	0.34	0.516	0.24	0.542	0.10	0.863	0.02	0.963	0.16	0.712	**1.53**	**0.022**	0.81	0.198	0.34	0.552
chr8:41519304	0.97	0.083	0.74	0.196	0.04	0.907	0.49	0.436	0.26	0.669	0.05	0.892	**1.51**	**0.034**	**1.42**	**0.034**	0.01	0.986
chr8:41519308	0.35	0.600	0.54	0.368	0.42	0.419	0.12	0.872	0.08	0.895	0.58	0.261	1.00	0.230	1.20	0.110	0.06	0.938
chr8:41519348	0.84	0.171	0.89	0.114	0.35	0.580	0.27	0.692	0.39	0.492	0.46	0.484	**1.55**	**0.034**	**1.58**	**0.025**	-0.02	0.985
chr8:41519399	**1.99**	**0.020**	0.68	0.388	0.31	0.743	1.19	0.208	0.70	0.399	0.25	0.791	**2.81**	**0.010**	0.45	0.645	0.31	0.791
chr8:41519411	**1.44**	**0.010**	0.03	0.970	0.12	0.829	0.76	0.228	0.17	0.845	-0.01	0.993	**2.19**	**0.001**	0.39	0.716	0.36	0.664
chr8:41519417	**1.28**	**0.032**	-0.22	0.758	0.03	0.946	0.37	0.588	0.55	0.468	-0.07	0.888	**2.44**	**0.001**	0.44	0.635	0.25	0.696
chr8:41519420	**1.95**	**0.003**	-0.75	0.302	-0.29	0.381	1.05	0.150	0.70	0.375	-0.41	0.272	**2.92**	**2.45E-04**	0.64	0.502	0.05	0.888
Amplicon Average	**1.08**	**0.024**	0.23	0.641	0.16	0.623	0.51	0.370	0.06	0.920	0.14	0.706	**1.99**	**0.001**	0.75	0.229	0.20	0.638

Supplementary Table 4: *ANK1* CpG sites are hypermethylated in the EC in individuals with VaD and co-existing AD pathology. Shown for each tissue is chromosomal location (genome build hg19) of CpG sites tested, with corrected difference (Δ) in DNA methylation levels and corresponding *P* value between controls and VaD cases after adjusting for the covariates of age, sex and batch effects. Analyses are stratified into "All VaD", "VaD Only" and "VaD + AD" groups. A positive Δ reflects hypermethylation in disease. All *P* values < 0.05 (and their corresponding Δ) are shown in bold.

	All VaD						VaD Only						VaD + AD					
	EC		STG		CER		EC		STG		CER		EC		STG		CER	
Genomic Location	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***
chr8:41519302	1.06	0.214	0.17	0.807	0.82	0.221	0.39	0.668	-0.24	0.742	0.72	0.322	**3.94**	**0.026**	1.87	0.180	1.11	0.429
chr8:41519304	1.38	0.127	0.16	0.827	0.52	0.404	0.66	0.492	-0.27	0.728	0.43	0.521	**4.46**	**0.020**	1.94	0.197	0.80	0.534
chr8:41519308	1.88	0.063	0.49	0.561	0.64	0.389	1.04	0.327	-0.03	0.975	0.82	0.314	**5.42**	**0.012**	2.61	0.140	0.13	0.931
chr8:41519348	0.78	0.412	-0.83	0.257	0.57	0.520	0.12	0.904	-1.29	0.092	0.44	0.647	3.62	0.063	1.23	0.406	1.03	0.577
chr8:41519399	1.47	0.278	1.49	0.176	1.96	0.132	0.52	0.716	1.36	0.233	1.56	0.261	**5.76**	**0.048**	1.79	0.437	3.40	0.206
chr8:41519411	1.53	0.068	**-2.66**	**0.026**	0.47	0.631	1.16	0.192	**-2.93**	**0.020**	0.18	0.867	3.02	0.105	-0.96	0.704	1.62	0.441
chr8:41519417	1.78	0.077	-0.84	0.418	0.69	0.351	0.89	0.394	-1.07	0.328	0.41	0.606	**5.72**	**0.008**	0.29	0.892	1.77	0.266
chr8:41519420	1.34	0.164	-1.29	0.217	-0.23	0.589	1.46	0.139	-1.62	0.146	-0.41	0.376	0.55	0.794	0.40	0.858	0.54	0.517
Amplicon Average	1.38	0.073	-0.39	0.546	0.73	0.122	0.82	0.325	-0.74	0.301	0.59	0.260	**4.06**	**0.016**	1.13	0.429	1.32	0.196

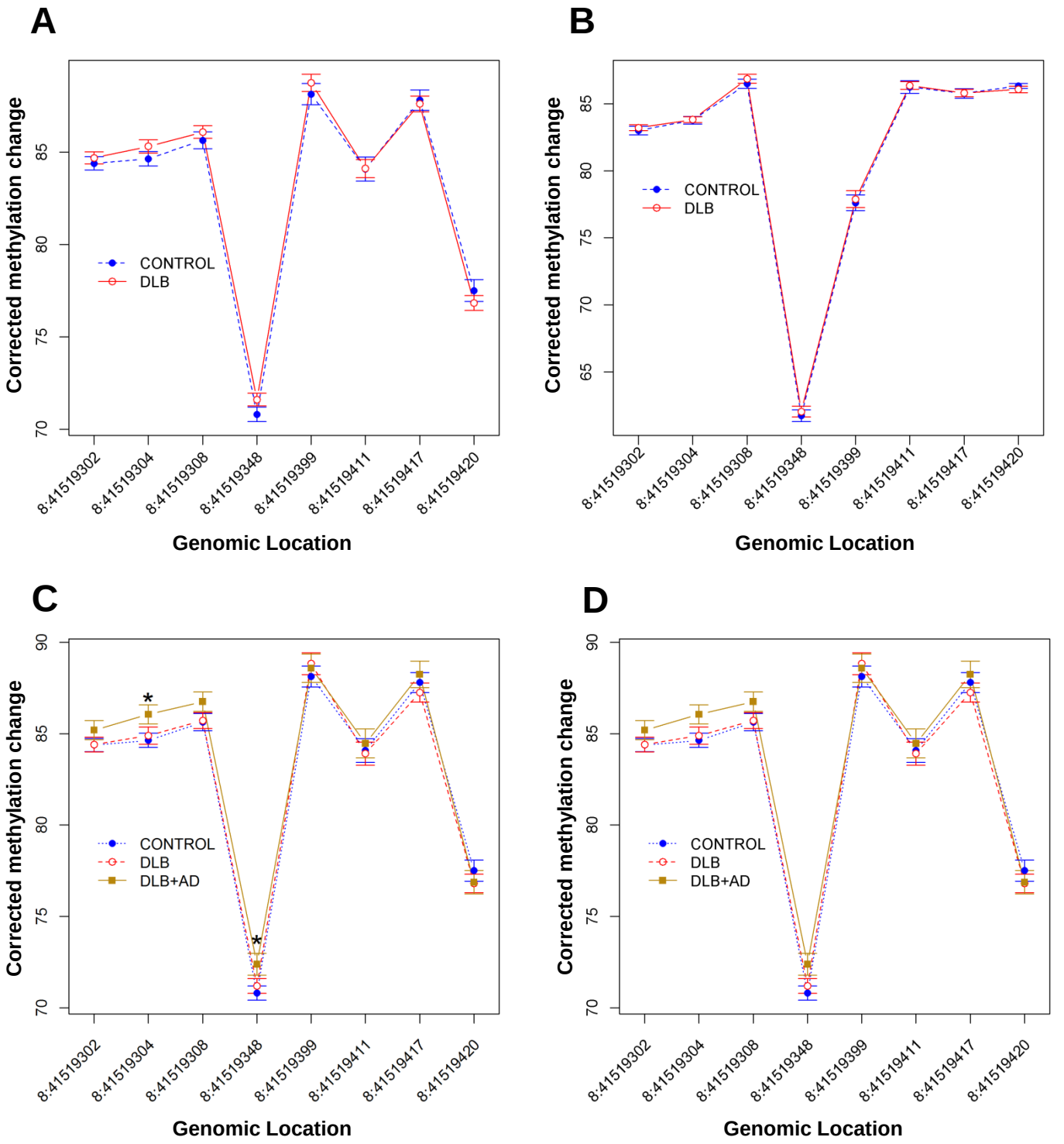
Supplementary Table 5: *ANK1* CpG sites are hypermethylated in the EC in HD. Shown for each tissue is chromosomal location (genome build hg19) of CpG sites tested, with corrected difference (Δ) in DNA methylation levels and corresponding *P* value between controls and HD cases after adjusting for the covariates of age, sex and batch effects. A positive Δ reflects hypermethylation in disease. All *P* values < 0.05 (and their corresponding Δ) are shown in bold.

	HD							
	EC		STG		CER		STR	
Genomic Location	Δ	*P*	Δ	*P*	Δ	*P*	Δ	*P*
chr8:41519302	**2.09**	**0.027**	1.20	0.160	0.33	0.681	0.76	0.496
chr8:41519304	1.89	0.062	0.58	0.532	0.31	0.666	1.14	0.377
chr8:41519308	**3.41**	**0.003**	1.61	0.136	0.99	0.261	-0.48	0.718
chr8:41519348	**2.54**	**0.014**	-0.24	0.797	-0.80	0.470	0.16	0.885
chr8:41519399	**3.89**	**0.014**	1.25	0.388	-0.68	0.673	-0.23	0.901
chr8:41519411	1.70	0.083	0.38	0.814	0.54	0.649	0.60	0.613
chr8:41519417	1.51	0.168	0.08	0.950	-0.05	0.960	0.82	0.461
chr8:41519420	**4.31**	**1.24E-04**	2.44	0.089	-0.37	0.509	1.63	0.184
Amplicon Average	**2.21**	**0.007**	0.63	0.431	0.07	0.903	0.84	0.356

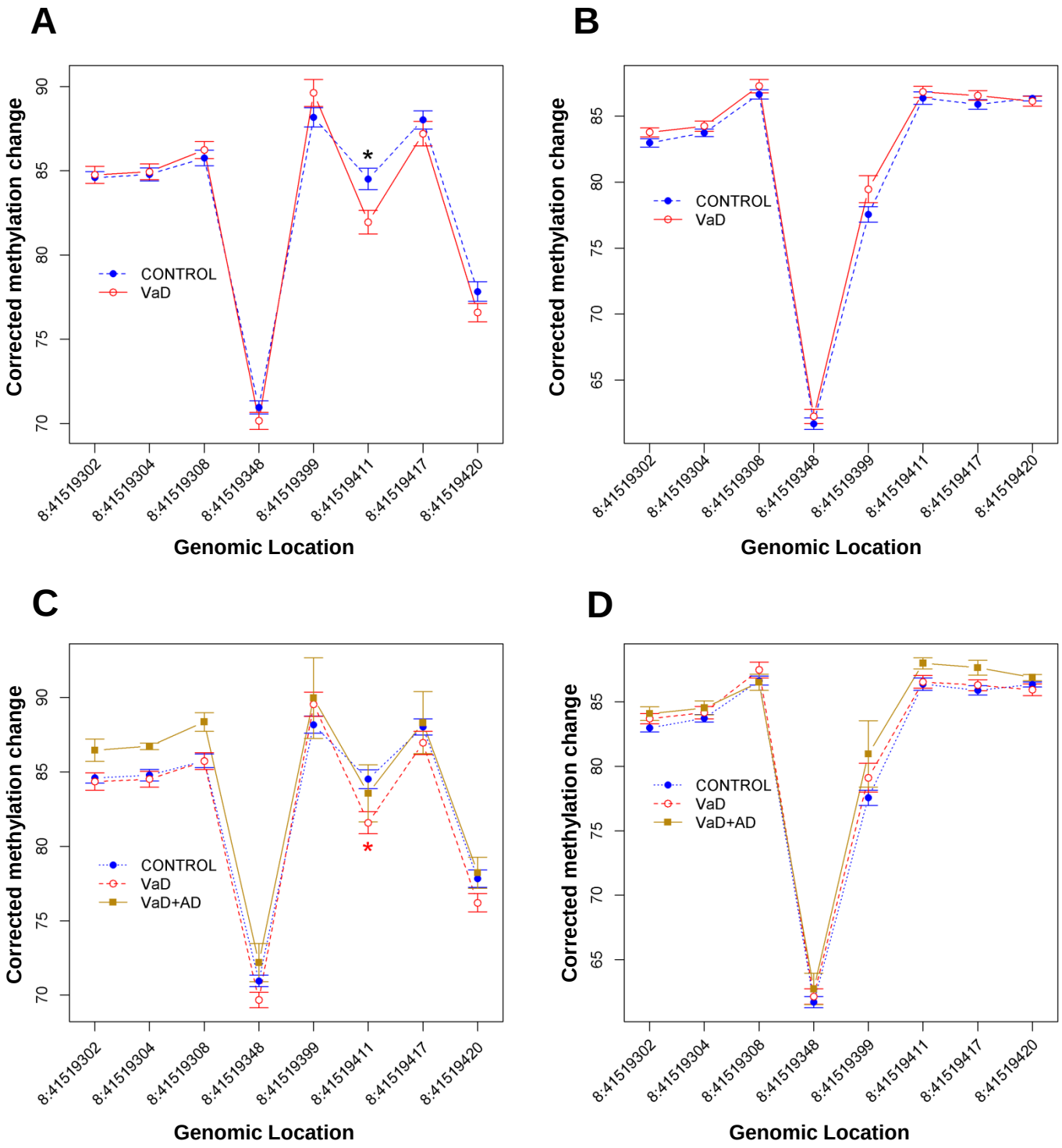
Supplementary Table 6: *ANK1* CpG sites are hypermethylated in the EC in PD. Shown for each tissue is chromosomal location (genome build hg19) of CpG sites tested, with corrected difference (Δ) in DNA methylation levels and corresponding *P* value between controls and PD cases after adjusting for the covariates of age, sex and batch effects. A positive Δ reflects hypermethylation in disease. All *P* values < 0.05 (and their corresponding Δ) are shown in bold.

	PD									
	EC		STG		CER		STR		SN	
Genomic Location	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***	Δ	***P***
chr8:41519302	0.37	0.592	0.20	0.777	0.23	0.714	1.43	0.351	-0.25	0.745
chr8:41519304	0.36	0.628	0.49	0.524	-0.33	0.562	0.89	0.647	0.19	0.823
chr8:41519308	0.34	0.686	0.80	0.334	-0.55	0.518	-2.46	0.203	0.05	0.956
chr8:41519348	0.83	0.323	0.07	0.932	0.18	0.838	1.17	0.381	-0.24	0.826
chr8:41519399	0.03	0.981	-0.10	0.924	0.16	0.893	-3.40	0.200	1.37	0.272
chr8:41519411	**1.65**	**0.032**	-1.32	0.234	0.14	0.875	-0.06	0.978	0.57	0.600
chr8:41519417	0.95	0.261	-0.87	0.371	-0.42	0.551	-0.98	0.521	0.06	0.950
chr8:41519420	**1.81**	**0.026**	-1.07	0.274	0.28	0.434	-0.01	0.997	0.36	0.714
Amplicon Average	0.84	0.202	-0.28	0.670	-0.02	0.959	-0.19	0.879	0.29	0.690

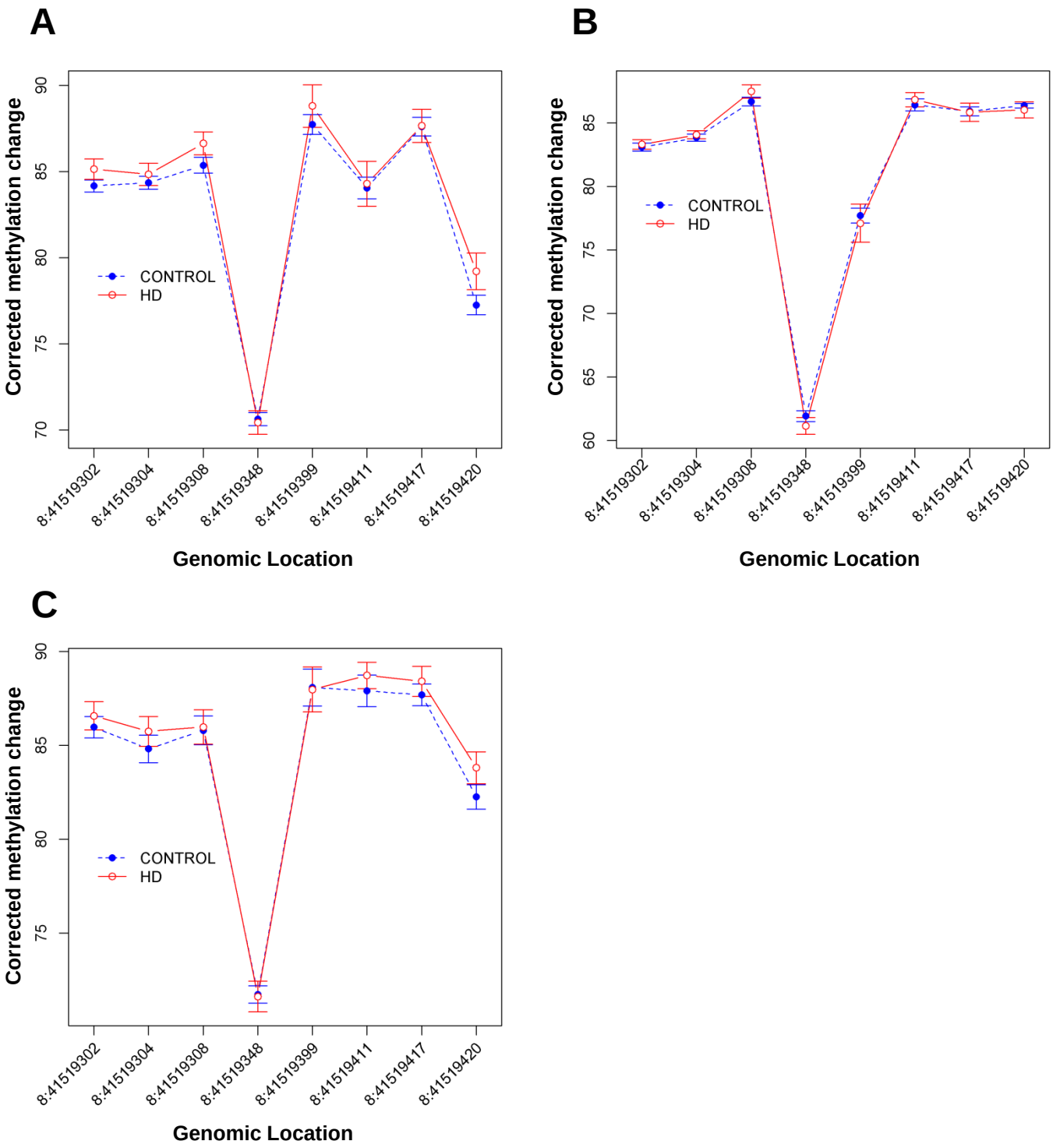
Supplementary Figure 1: *ANK1* DNA methylation levels in the STG and CER in individuals with DLB Using bisulfite pyrosequencing we assayed a 118bp region of the *ANK1* gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in all DLB samples compared to control samples. We did not observe any disease-associated changes in either the STG (**A**) or CER (**B**). When we compared DNA methylation levels in *ANK1* in individuals with DLB and co-existing AD pathology, or individuals with "pure" DLB only to controls, we found significant hypermethylation at two sites in the STG (**C**) but no sites in the CER (**D**) in individuals with co-existing AD pathology. Data is represented as mean (±SEM) Key: **P* < 0.05, ***P* < 0.01, ****P* < 0.005.



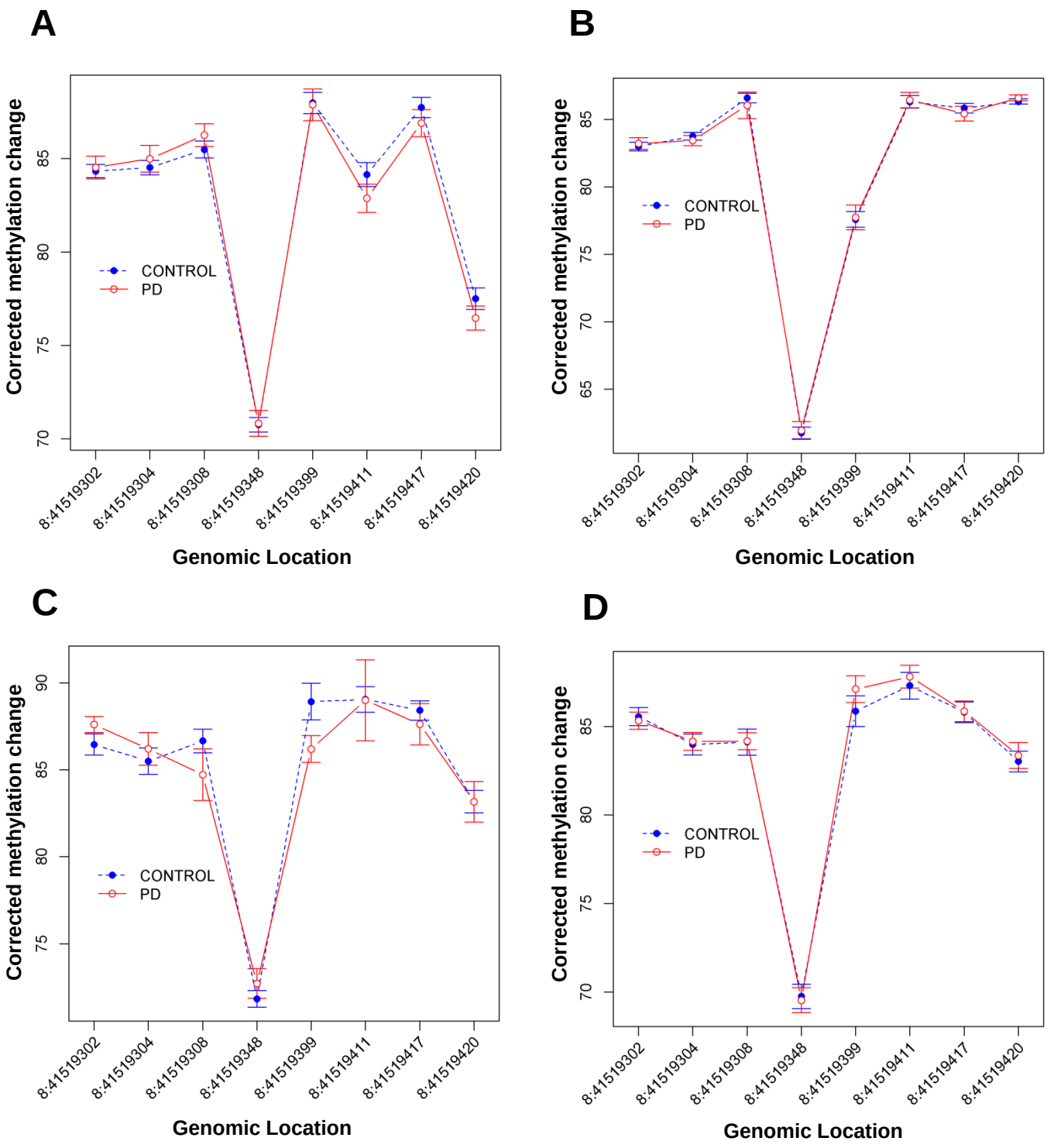
Supplementary Figure 2: *ANK*1 DNA methylation levels in the STG and CER in individuals with VaD. Using bisulfite pyrosequencing we assayed a 118bp region of the *ANK1* gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in all VaD samples compared to control samples. We observed neuropathology-associated hypomethylation at one site in the STG (**A**) and no disease-associated changes in the CER (**B**). When we compared DNA methylation levels in *ANK1* in individuals with VaD and co-existing AD pathology, or individuals with "pure" VaD only to controls, we found significant hypomethylation at one sites in the STG in individuals with "pure" VaD (**C**), but no disease-associated differences in the CER (**D**). Data is represented as mean (±SEM) Key: **P* < 0.05, ***P* < 0.01, ****P* < 0.005.



Supplementary Figure 3 *ANK*1 DNA methylation levels in the STG, CER and STR in individuals with HD Using bisulfite pyrosequencing we assayed a 118bp region of the *ANK1* gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in HD samples compared to control samples. We did not observe any disease-associated changes in either the STG (**A**), CER (**B**) or STR (**C**). Data is represented as mean (±SEM) Key: **P* < 0.05, ***P* < 0.01, ****P* < 0.005.



Supplementary Figure 4 *ANK*1 DNA methylation levels in the STG, CER, STR and SN in individuals with PD Using bisulfite pyrosequencing we assayed a 118bp region of the *ANK1* gene ranging from 41519302 to 41519420 on chromosome 8 (genome build hg19) in PD samples compared to control samples. We did not observe any disease-associated changes in either the STG (**A**), CER (**B**), STR (**C**) or SN **(D).** Data is represented as mean (±SEM) Key: **P* < 0.05, ***P* < 0.01, ****P* < 0.005.



BIBLIOGRAPHY

Abraham, R., V. Moskvina, R. Sims, P. Hollingworth, A. Morgan, L. Georgieva, K. Dowzell, S. Cichon, A. M. Hillmer, M. C. O'Donovan, J. Williams, M. J. Owen and G. Kirov (2008). "A genome-wide association study for late-onset Alzheimer's disease using DNA pooling." BMC Med Genomics **1**: 44.

Adrain, C., M. Zettl, Y. Christova, N. Taylor and M. Freeman (2012). "Tumor Necrosis Factor Signaling Requires iRhom2 to Promote Trafficking and Activation of TACE." Science **335**(6065): 225-228.

Adriana Martorana, Matteo Bulati, Silvio Buffa, Mariavaleria Pellicanò, Calogero Caruso, G. Candore and G. Colonna-Romano (2012). "Immunosenescence, inflammation and Alzheimer's disease." Longevity and Lifespan **1**.

Agbemenyah, H. Y., R. C. Agis-Balboa, S. Burkhardt, I. Delalle and A. Fischer (2014). "Insulin growth factor binding protein 7 is a novel target to treat dementia." Neurobiol Dis **62**: 135-143.

Aran, D., G. Toperoff, M. Rosenberg and A. Hellman (2011). "Replication timing-related and gene body-specific methylation of active human genes." Hum Mol Genet **20**(4): 670-680.

Arispe, N. and M. Doh (2002). "Plasma membrane cholesterol controls the cytotoxicity of Alzheimer's disease AβP (1–40) and (1–42) peptides." The FASEB Journal **16**(12): 1526-1536.

Arosio, B., A. Bulbarelli, S. Bastias Candia, E. Lonati, L. Mastronardi, P. Romualdi, S. Candeletti, C. Gussago, D. Galimberti, E. Scarpini, B. Dell'Osso, C. Altamura, M. MacCarrone, L. Bergamaschini, C. D'Addario and D. Mari (2012). "Pin1 contribution to Alzheimer's disease: transcriptional and epigenetic mechanisms in patients with late-onset Alzheimer's disease." Neurodegener Dis **10**(1-4): 207-211.

Arvanitakis, Z., R. S. Wilson, J. L. Bienias, D. A. Evans and D. A. Bennett (2004). "Diabetes mellitus and risk of alzheimer disease and decline in cognitive function." Archives of Neurology **61**(5): 661-666.

Azuara, V., P. Perry, S. Sauer, M. Spivakov, H. F. Jorgensen, R. M. John, M. Gouti, M. Casanova, G. Warnes, M. Merkenschlager and A. G. Fisher (2006). "Chromatin signatures of pluripotent cell lines." Nat Cell Biol **8**(5): 532-538.

Bachman, M., S. Uribe-Lewis, X. Yang, M. Williams, A. Murrell and S. Balasubramanian (2014). "5-Hydroxymethylcytosine is a predominantly stable DNA modification." Nature chemistry **6**.

Bakulski, K. M., D. C. Dolinoy, M. A. Sartor, H. L. Paulson, J. R. Konen, A. P. Lieberman, R. L. Albin, H. Hu and L. S. Rozek (2012). "Genome-wide DNA methylation differences between late-onset Alzheimer's disease and cognitively normal controls in human frontal cortex." Journal of Alzheimers Disease **29**(3): 571-588.

Bakulski, K. M., D. C. Dolinoy, M. A. Sartor, H. L. Paulson, J. R. Konen, A. P. Lieberman, R. L. Albin, H. Hu and L. S. Rozek (2012). "Genome-wide DNA methylation differences between late-onset Alzheimer's disease and cognitively normal controls in human frontal cortex." J Alzheimers Dis **29**(3): 571-588.

Balasubramanian, D., B. Akhtar-Zaidi, L. Song, C. F. Bartels, M. Veigl, L. Beard, L. Myeroff, K. Guda, J. Lutterbaugh, J. Willis, G. E. Crawford, S. D. Markowitz and P. C. Scacheri (2012). "H3K4me3 inversely correlates with DNA methylation at a large class of non-CpG-island-containing start sites." Genome Med **4**(5): 47.

Ball, M. P., J. B. Li, Y. Gao, J. H. Lee, E. M. LeProust, I. H. Park, B. Xie, G. Q. Daley and G. M. Church (2009). "Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells." Nat Biotechnol **27**(4): 361-368.

Barger, S. W. and A. S. Basile (2001). "Activation of microglia by secreted amyloid precursor protein evokes release of glutamate by cystine exchange and attenuates synaptic function." Journal of Neurochemistry **76**(3): 846-854.

Barrachina, M. and I. Ferrer (2009). "DNA methylation of Alzheimer disease and tauopathy-related genes in postmortem brain." J Neuropathol Exp Neurol **68**(8): 880-891.

Barski, A., S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev and K. Zhao (2007). "High-resolution profiling of histone methylations in the human genome." Cell **129**(4): 823-837.

Bartel, D. P. (2004). "MicroRNAs: Genomics, Biogenesis, Mechanism, and Function." Cell **116**(2): 281-297.

Barth, T. K. and A. Imhof (2010). "Fast signals and slow marks: the dynamics of histone modifications." Trends Biochem Sci **35**(11): 618-626.

Behl, C. (2000). "Apoptosis and Alzheimer's disease." J Neural Transm **107**(11): 1325-1344.

Bekris, L. M., C.-E. Yu, T. D. Bird and D. W. Tsuang (2010). "Review Article: Genetics of Alzheimer Disease." Journal of Geriatric Psychiatry and Neurology **23**(4): 213-227.

Bernstein, B. E., M. Kamal, K. Lindblad-Toh, S. Bekiranov, D. K. Bailey, D. J. Huebert, S. McMahon, E. K. Karlsson, E. J. Kulbokas, 3rd, T. R. Gingeras, S. L. Schreiber and E. S. Lander (2005). "Genomic maps and comparative analysis of histone modifications in human and mouse." Cell **120**(2): 169-181.

Bernstein, B. E., T. S. Mikkelsen, X. Xie, M. Kamal, D. J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S. L. Schreiber and E. S. Lander (2006). "A bivalent chromatin structure marks key developmental genes in embryonic stem cells." Cell **125**(2): 315-326.

Bibikova, M., B. Barnes, C. Tsan, V. Ho, B. Klotzle, J. M. Le, D. Delano, L. Zhang, G. P. Schroth, K. L. Gunderson, J. B. Fan and R. Shen (2011). "High density DNA methylation array with single CpG site resolution." Genomics **98**.

Bird, A. (2007). "Perceptions of epigenetics." Nature **447**.

Bollati, V., D. Galimberti, L. Pergoli, E. Dalla Valle, F. Barretta, F. Cortini, E. Scarpini, P. A. Bertazzi and A. Baccarelli (2011). "DNA methylation in repetitive elements and Alzheimer disease." Brain Behav Immun **25**(6): 1078-1083.

Booth, M. J., M. R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik and S. Balasubramanian (2012). "Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution." Science **336**(6083): 934-937.

Booth, M. J., T. W. B. Ost, D. Beraldi, N. M. Bell, M. R. Branco, W. Reik and S. Balasubramanian (2013). "Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine." Nat. Protocols **8**(10): 1841-1851.

Boyer, L. A., K. Plath, J. Zeitlinger, T. Brambrink, L. A. Medeiros, T. I. Lee, S. S. Levine, M. Wernig, A. Tajonar, M. K. Ray, G. W. Bell, A. P. Otte, M. Vidal, D. K. Gifford, R. A. Young and R. Jaenisch (2006). "Polycomb complexes repress developmental regulators in murine embryonic stem cells." Nature **441**(7091): 349-353.

Braak, H. and E. Braak (1991). "Neuropathological stageing of Alzheimer-related changes." Acta Neuropathol **82**(4): 239-259.

Braak, H., E. Braak, J. Bohl and W. Lang (1989). "Alzheimer's disease: amyloid plaques in the cerebellum." Journal of the Neurological Sciences **93**(2): 277-287.

Bradley-Whitman, M. A. and M. A. Lovell (2013). "Epigenetic changes in the progression of Alzheimer's disease." Mech Ageing Dev **134**(10): 486-495.

Branco, M. R., G. Ficz and W. Reik (2012). "Uncovering the role of 5-hydroxymethylcytosine in the epigenome." Nat Rev Genet **13**(1): 7-13.

Breiling, A. and F. Lyko (2015). "Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond." Epigenetics & Chromatin **8**(1): 24.

Breslow, J. L., V. I. Zannis, T. R. SanGiacomo, J. L. Third, T. Tracy and C. J. Glueck (1982). "Studies of familial type III hyperlipoproteinemia using as a genetic marker the apoE phenotype E2/2." Journal of Lipid Research **23**(8): 1224-1235.

Brohede, J., M. Rinde, B. Winblad and C. Graff (2010). "A DNA methylation study of the amyloid precursor protein gene in several brain regions from patients with familial Alzheimer disease." J Neurogenet **24**(4): 179-181.

Brunnström, H. R. and E. M. Englund (2009). "Cause of death in patients with dementia disorders." European Journal of Neurology **16**(4): 488-492.
Burns, A. and S. Iliffe (2009). "Alzheimer's disease." BMJ **338**.
Caricasole, A., A. Copani, F. Caraci, E. Aronica, A. J. Rozemuller, A. Caruso, M. Storto, G. Gaviraghi, G. C. Terstappen and F. Nicoletti (2004). "Induction of Dickkopf-1, a negative modulator of the Wnt pathway, is associated with neuronal degeneration in Alzheimer's brain." J Neurosci **24**(26): 6021-6027.
Carter, D. B. (2005). "The interaction of amyloid-beta with ApoE." Subcell Biochem **38**: 255-272.
Cedar, H. and Y. Bergman (2009). "Linking DNA methylation and histone modification: patterns and paradigms." Nat Rev Genet **10**(5): 295-304.
Celarain, N., J. Sanchez-Ruiz de Gordoa, M. V. Zelaya, M. Roldan, R. Larumbe, L. Pulido, C. Echavarri and M. Mendioroz (2016). "TREM2 upregulation correlates with 5-hydroxymethycytosine enrichment in Alzheimer's disease hippocampus." Clin Epigenetics **8**: 37.
Champely, S., C. Ekstrom, P. Dalgaard, J. Gill, S. Weibelzahl, A. Anandkumar, C. Ford, R. Volcic and H. De Rosario (2018). "Basic Functions for Power Analysis." CRAN **1.2-2**.
Chan, G., C. C. White, P. A. Winn, M. Cimpean, J. M. Replogle, L. R. Glick, N. E. Cuerdon, K. J. Ryan, K. A. Johnson, J. A. Schneider, D. A. Bennett, L. B. Chibnik, R. A. Sperling, E. M. Bradshaw and P. L. De Jager (2015). "CD33 modulates TREM2: convergence of Alzheimer loci." Nat Neurosci **18**(11): 1556-1558.
Chen, Y.-a., M. Lemire, S. Choufani, D. T. Butcher, D. Grafodatskaya, B. W. Zanke, S. Gallinger, T. J. Hudson and R. Weksberg (2013). "Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray." Epigenetics **8**(2): 203-209.
Chen, Y. A., M. Lemire, S. Choufani, D. T. Butcher, D. Grafodatskaya, B. W. Zanke, S. Gallinger, T. J. Hudson and R. Weksberg (2013). "Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray." Epigenetics **8**(2): 203-209.
Chi, P., C. D. Allis and G. G. Wang (2010). "Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers." Nat Rev Cancer **10**(7): 457-469.
Chi, S., J. H. Song, M. S. Tan, W. Zhang, Z. X. Wang, T. Jiang, L. Tan and J. T. Yu (2015). "Association of Single-Nucleotide Polymorphism in ANK1 with Late-Onset Alzheimer's Disease in Han Chinese." Mol Neurobiol.
Chibnik, L. B., L. Yu, M. L. Eaton, G. Srivastava, J. A. Schneider, M. Kellis, D. A. Bennett and P. L. De Jager (2015). "Alzheimer's loci: epigenetic associations and interaction with genetic factors." Ann Clin Transl Neurol **2**(6): 636-647.
Choi, J. H., C. H. Lee, I. K. Hwang, M. H. Won, J. K. Seong, Y. S. Yoon, H. S. Lee and I. S. Lee (2007). "Age-related changes in ionized calcium-binding adapter molecule 1 immunoreactivity and protein level in the gerbil hippocampal CA1 region." J Vet Med Sci **69**(11): 1131-1136.
Chouliaras, L., D. Mastroeni, E. Delvaux, A. Grover, G. Kenis, P. R. Hof, H. W. Steinbusch, P. D. Coleman, B. P. Rutten and D. L. van den Hove (2013). "Consistent decrease in global DNA methylation and hydroxymethylation in the hippocampus of Alzheimer's disease patients." Neurobioly of Aging **34**(9): 2091-2099.
Chung, W.-S., P. B. Verghese, C. Chakraborty, J. Joung, B. T. Hyman, J. D. Ulrich, D. M. Holtzman and B. A. Barres (2016). "Novel allele-dependent role for APOE in controlling the rate of synapse pruning by astrocytes." Proceedings of the National Academy of Sciences **113**(36): 10186-10191.
Clouaire, T., S. Webb, P. Skene, R. Illingworth, A. Kerr, R. Andrews, J. H. Lee, D. Skalnik and A. Bird (2012). "Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells." Genes Dev **26**(15): 1714-1728.
Collas, P. (2010). "The Current State of Chromatin Immunoprecipitation." Molecular Biotechnology **45**(1): 87-100.
Condliffe, D., A. Wong, C. Troakes, P. Proitsi, Y. Patel, L. Chouliaras, C. Fernandes, J. Cooper, S. Lovestone, L. Schalkwyk, J. Mill and K. Lunnon (2014). "Cross-region reduction in 5-

hydroxymethylcytosine in Alzheimer's disease brain." Neurobiology of Aging **35**(8): 1850-1854.
Condliffe, D., A. Wong, C. Troakes, P. Proitsi, Y. Patel, L. Chouliaras, C. Fernandes, J. Cooper, S. Lovestone, L. Schalkwyk, J. Mill and K. Lunnon (2014). "Cross-region reduction in 5-hydroxymethylcytosine in Alzheimer's disease brain." Neurobiol Aging **35**(8): 1850-1854.
Cong, L., J. Jia, W. Qin, Y. Ren and Y. Sun (2014). "Genome-wide analysis of DNA methylation in an APP/PS1 mouse model of Alzheimer's disease." Acta Neurol Belg **114**(3): 195-206.
Coon.;, K. D., A. J. Myers.;, D. W. Craig.;, J. A. Webste.;, J. V. Pearson.;, D. H. Lince.;, V. L. Zismann.;, T. G. Beach.;, D. Leung.;, L. Bryden.;, R. F. Halperin.;, L. Marlowe.;, M. Kaleem.;, D. G. Walker.;, R. Ravid.;, C. B. Heward.;, J. Rogers.;, A. Papassotiropoulos.;, E. M. Reiman.;, J. Hardy.; and D. A. Stephan. (2007). "A High-Density Whole-Genome Association Study Reveals That APOE Is the Major Susceptibility Gene for Sporadic Late-Onset Alzheimer's Disease." Journal of Clinical Psychiatry **68**(4): 613-618.
Coppieters, N., B. V. Dieriks, C. Lill, R. L. Faull, M. A. Curtis and M. Dragunow (2014). "Global changes in DNA methylation and hydroxymethylation in Alzheimer's disease human brain." Neurobiol Aging **35**(6): 1334-1344.
Corder, E., A. Saunders, W. Strittmatter, D. Schmechel, P. Gaskell, G. Small, A. Roses, J. Haines and M. Pericak-Vance (1993). "Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families." Science **261**(5123): 921-923.
Corder, E. H., A. M. Saunders, N. J. Risch, W. J. Strittmatter, D. E. Schmechel, P. C. Gaskell, Jr., J. B. Rimmler, P. A. Locke, P. M. Conneally and K. E. Schmader (1994). "Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease." Nat Genet **7**(2): 180-184.
Cousin, M. A. and P. J. Robinson (2001). "The dephosphins: dephosphorylation by calcineurin triggers synaptic vesicle endocytosis." Trends Neurosci **24**(11): 659-665.
Cruchaga, C., J. S. Kauwe, P. Nowotny, K. Bales, E. H. Pickering, K. Mayo, S. Bertelsen, A. Hinrichs, A. M. Fagan, D. M. Holtzman, J. C. Morris and A. M. Goate (2012). "Cerebrospinal fluid APOE levels: an endophenotype for genetic studies for Alzheimer's disease." Hum Mol Genet **21**(20): 4558-4571.
Cruchaga, C., P. Nowotny, J. S. Kauwe, P. G. Ridge, K. Mayo, S. Bertelsen, A. Hinrichs, A. M. Fagan, D. M. Holtzman, J. C. Morris and A. M. Goate (2011). "Association and expression analyses with single-nucleotide polymorphisms in TOMM40 in Alzheimer disease." Arch Neurol **68**(8): 1013-1019.
Cummings, J. L. (2004). "Dementia with Lewy Bodies: Molecular Pathogenesis and Implications for Classification." Journal of Geriatric Psychiatry and Neurology **17**(3): 112-119.
Curradi, M., A. Izzo, G. Badaracco and N. Landsberger (2002). "Molecular Mechanisms of Gene Silencing Mediated by DNA Methylation." Molecular and Cellular Biology **22**(9): 3157-3173.
Daniel Lee, C. Y. and G. E. Landreth (2010). "The role of microglia in amyloid clearance from the AD brain." Journal of neural transmission (Vienna, Austria : 1996) **117**(8): 949-960.
Davies, C. A., D. M. A. Mann, P. Q. Sumpter and P. O. Yates (1987). "A quantitative morphometric analysis of the neuronal and synaptic content of the frontal and temporal cortex in patients with Alzheimer's disease." Journal of the Neurological Sciences **78**(2): 151-164.
Davis, S., P. Du, S. Bilke, J. Triche, =. and M. Bootwalla (2012). "Methylumi: Handle Illumina Methylation Data 2012..

." R package version 2.2.0.
De Jager, P. L., G. Srivastava, K. Lunnon, J. Burgess, L. C. Schalkwyk, L. Yu, M. L. Eaton, B. T. Keenan, J. Ernst, C. McCabe, A. Tang, T. Raj, J. Replogle, W. Brodeur, S. Gabriel, H. S. Chai, C. Younkin, S. G. Younkin, F. Zou, M. Szyf, C. B. Epstein, J. A. Schneider, B. E. Bernstein, A. Meissner, N. Ertekin-Taner, L. B. Chibnik, M. Kellis, J. Mill and D. A. Bennett (2014). "Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci." Nat Neurosci **17**(9): 1156-1163.

De Jager, P. L., G. Srivastava, K. Lunnon, J. Burgess, L. C. Schalkwyk, L. Yu, M. L. Eaton, B. T. Keenan, J. Ernst, C. McCabe, A. Tang, T. Raj, J. Replogle, W. Brodeur, S. Gabriel, H. S. Chai, C. Younkin, S. G. Younkin, F. Zou, M. Szyf, C. B. Epstein, J. A. Schneider, B. E. Bernstein, A. Meissner, N. Ertekin-Taner, L. B. Chibnik, M. Kellis, J. Mill and D. A. Bennett (2014). "Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci." Nat Neurosci **Sep;17**(9): 1156-1163.

Deaton, A. M. and A. Bird (2011). "CpG islands and the regulation of transcription." Genes & Development **25**(10): 1010-1022.

Dedeurwaerder, S., M. Defrance, E. Calonne, H. Denis, C. Sotiriou and F. Fuks (2011). "Evaluation of the Infinium Methylation 450K technology." Epigenomics **3**(6): 771-784.

Deissler, H., M. Wilm, B. Genc, B. Schmitz, T. Ternes, F. Naumann, M. Mann and W. Doerfler (1997). "Rapid protein sequencing by tandem mass spectrometry and cDNA cloning of p20-CGGBP. A novel protein that binds to the unstable triplet repeat 5'-d(CGG)n-3' in the human FMR1 gene." J Biol Chem **272**(27): 16761-16768.

Desikan, R. S., A. J. Schork, Y. Wang, A. Witoelar, M. Sharma, L. K. McEvoy, D. Holland, J. B. Brewer, C. H. Chen, W. K. Thompson, D. Harold, J. Williams, M. J. Owen, M. C. O'Donovan, M. A. Pericak-Vance, R. Mayeux, J. L. Haines, L. A. Farrer, G. D. Schellenberg, P. Heutink, A. B. Singleton, A. Brice, N. W. Wood, J. Hardy, M. Martinez, S. H. Choi, A. DeStefano, M. A. Ikram, J. C. Bis, A. Smith, A. L. Fitzpatrick, L. Launer, C. van Duijn, S. Seshadri, I. D. Ulstein, D. Aarsland, T. Fladby, S. Djurovic, B. T. Hyman, J. Snaedal, H. Stefansson, K. Stefansson, T. Gasser, O. A. Andreassen and A. M. Dale (2015). "Genetic overlap between Alzheimer's disease and Parkinson's disease at the MAPT locus." Mol Psychiatry **20**(12): 1588-1595.

Devall, M., J. Mill and K. Lunnon (2014). "The mitochondrial epigenome: a role in Alzheimer's disease?" Epigenomics **6**(6): 665-675.

Devall, M., J. Roubroeks, J. Mill, M. Weedon and K. Lunnon (2016). "Epigenetic regulation of mitochondrial function in neurodegenerative disease: New insights from advances in genomic technologies." Neurosci Lett.

Di Francesco, A., B. Arosio, C. Gussago, E. Dainese, D. Mari, C. D'Addario and M. Maccarrone (2013). "Involvement of 5-lipoxygenase in Alzheimer's disease: a role for DNA methylation." J Alzheimers Dis **37**(1): 3-8.

Di Paolo, G., S. Sankaranarayanan, M. R. Wenk, L. Daniell, E. Perucco, B. J. Caldarone, R. Flavell, M. R. Picciotto, T. A. Ryan, O. Cremona and P. De Camilli (2002). "Decreased synaptic vesicle recycling efficiency and cognitive deficits in amphiphysin 1 knockout mice." Neuron **33**(5): 789-804.

Dodel, R., I. Csoti, G. Ebersbach, G. Fuchs, M. Hahne, W. Kuhn, M. Oechsner, W. Jost, H. Reichmann and J. B. Schulz (2008). "Lewy body dementia and Parkinson's disease with dementia." J Neurol **255 Suppl 5**: 39-47.

Dreyling, M. H., J. A. Martinez-Climent, M. Zheng, J. Mao, J. D. Rowley and S. K. Bohlander (1996). "The t(10;11)(p13;q14) in the U937 cell line results in the fusion of the AF10 gene and CALM, encoding a new member of the AP-3 clathrin assembly protein family." Proc Natl Acad Sci U S A **93**(10): 4804-4809.

Du, P., X. Zhang, C. C. Huang, N. Jafari, W. A. Kibbe, L. Hou and S. M. Lin (2010). "Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis." BMC Bioinformatics **11**.

Elias-Sonnenschein, L. S., S. Helisalmi, T. Natunen, A. Hall, T. Paajanen, S. K. Herukka, M. Laitinen, A. M. Remes, A. M. Koivisto, K. M. Mattila, T. Lehtimaki, F. R. Verhey, P. J. Visser, H. Soininen and M. Hiltunen (2013). "Genetic loci associated with Alzheimer's disease and cerebrospinal fluid biomarkers in a Finnish case-control cohort." PLoS One **8**(4): e59676.

Ellison, E. M., M. A. Bradley-Whitman and M. A. Lovell (2017). "Single-Base Resolution Mapping of 5-Hydroxymethylcytosine Modifications in Hippocampus of Alzheimer's Disease Subjects." J Mol Neurosci.

Escott-Price, V., R. Sims, C. Bannister, D. Harold, M. Vronskaya, E. Majounie, N. Badarinarayan, Gerad/Perades, I. consortia, K. Morgan, P. Passmore, C. Holmes, J. Powell, C. Brayne, M. Gill, S. Mead, A. Goate, C. Cruchaga, J. C. Lambert, C. van Duijn, W. Maier, A. Ramirez, P. Holmans, L. Jones, J. Hardy, S. Seshadri, G. D. Schellenberg, P. Amouyel and J. Williams (2015). "Common polygenic variation enhances risk prediction for Alzheimer's disease." Brain **138**(Pt 12): 3673-3684.

Esiri, M. M. (1993). "Brain banks: the Oxford experience." J Neural Transm Suppl **39**: 25-30.

Esteller, M. (2008). "Epigenetics in Cancer." New England Journal of Medicine **358**(11): 1148-1159.

Farooqui, A., K. Wells and L. Horrocks (1995). "Breakdown of membrane phospholipids in Alzheimer disease." Molecular and Chemical Neuropathology **25**(2-3): 155-173.

Fassbender, K., M. Simons, C. Bergmann, M. Stroick, D. Lütjohann, P. Keller, H. Runz, S. Kühl, T. Bertsch, K. von Bergmann, M. Hennerici, K. Beyreuther and T. Hartmann (2001). "Simvastatin strongly reduces levels of Alzheimer's disease β-amyloid peptides Aβ42 and Aβ40 in vitro and in vivo." Proceedings of the National Academy of Sciences **98**(10): 5856-5861.

Fearnley, J. M. and A. J. Lees (1991). "Aging and Parkinson's disease: substantia nigra regional selectivity." Brain **114**(5): 2283-2301.

Feinberg, A. P. and B. Tycko (2004). "The history of cancer epigenetics." Nat Rev Cancer **4**(2): 143-153.

Feulner, T. M., S. M. Laws, P. Friedrich, S. Wagenpfeil, S. H. Wurst, C. Riehle, K. A. Kuhn, M. Krawczak, S. Schreiber, S. Nikolaus, H. Forstl, A. Kurz and M. Riemenschneider (2010). "Examination of the current top candidate genes for AD in a genome-wide association study." Mol Psychiatry **15**(7): 756-766.

Field, S. F., D. Beraldi, M. Bachman, S. K. Stewart, S. Beck and S. Balasubramanian (2015). "Accurate measurement of 5-methylcytosine and 5-hydroxymethylcytosine in human cerebellum DNA by oxidative bisulfite on an array (OxBS-array)." PLoS One **10**(2): e0118202.

Fletcher, D. A. and R. D. Mullins (2010). "Cell mechanics and the cytoskeleton." Nature **463**(7280): 485-492.

Flores, K., F. Wolschin, J. J. Corneveaux, A. N. Allen, M. J. Huentelman and G. V. Amdam (2012). "Genome-wide association between DNA methylation and alternative splicing in an invertebrate." BMC Genomics **13**: 480.

Flusberg, B. A., D. R. Webster, J. H. Lee, K. J. Travers, E. C. Olivares, T. A. Clark, J. Korlach and S. W. Turner (2010). "Direct detection of DNA methylation during single-molecule, real-time sequencing." Nat Meth **7**(6): 461-465.

Forabosco, P., A. Ramasamy, D. Trabzuni, R. Walker, C. Smith, J. Bras, A. P. Levine, J. Hardy, J. M. Pocock, R. Guerreiro, M. E. Weale and M. Ryten (2013). "Insights into TREM2 biology by network analysis of human brain gene expression data." Neurobiol Aging **34**(12): 2699-2714.

Francis, Y. I., M. Fà, H. Ashraf, H. Zhang, A. Staniszewski, D. S. Latchman and O. Arancio (2009). "Dysregulation of Histone Acetylation in the APP/PS1 Mouse Model of Alzheimer's Disease." Journal of Alzheimer's Disease **18**(1): 131-139.

Fratiglioni, L., L. Launer, K. Andersen, M. Breteler, J. Copeland, J. Dartigues, A. Lobo, J. Martinez-Lage, H. Soininen and A. Hofman (1999). "Incidence of dementia and major subtypes in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group." Neurology **54**(11 Suppl 5): S10-15.

Frauer, C., T. Hoffmann, S. Bultmann, V. Casa, M. C. Cardoso, I. Antes and H. Leonhardt (2011). "Recognition of 5-Hydroxymethylcytosine by the Uhrf1 SRA Domain." PLOS ONE **6**(6): e21306.

Frommer, M., L. E. McDonald, D. S. Millar, C. M. Collis, F. Watt, G. W. Grigg, P. L. Molloy and C. L. Paul (1992). "A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands." Proc Natl Acad Sci U S A **89**(5): 1827-1831.

Furuya, T. K., P. N. da Silva, S. L. Payao, L. T. Rasmussen, R. W. de Labio, P. H. Bertolucci, I. L. Braga, E. S. Chen, G. Turecki, N. Mechawar, J. Mill and M. de Arruda Cardoso Smith (2012). "SORL1 and SIRT1 mRNA expression and promoter methylation levels in aging and Alzheimer's Disease." Neurochem Int **61**(7): 973-975.

Furuya, T. K., P. N. Silva, S. L. Payao, P. H. Bertolucci, L. T. Rasmussen, R. W. De Labio, I. L. Braga, E. S. Chen, G. Turecki, N. Mechawar, J. Mill and M. A. Smith (2012). "Analysis of SNAP25 mRNA expression and promoter DNA methylation in brain areas of Alzheimer's Disease patients." Neuroscience **220**: 41-46.

Gallagher, P. G. and B. G. Forget (1998). "An alternate promoter directs expression of a truncated, muscle-specific isoform of the human ankyrin 1 gene." J Biol Chem **273**(3): 1339-1348.

Gallagher, P. G., W. T. Tse, A. L. Scarpa, S. E. Lux and B. G. Forget (1997). "Structure and Organization of the Human Ankyrin-1 Gene: BASIS FOR COMPLEXITY OF PRE-mRNA PROCESSING." Journal of Biological Chemistry **272**(31): 19220-19228.

Gatz, M., C. A. Reynolds, L. Fratiglioni, B. Johansson, J. A. Mortimer, S. Berg, A. Fiske and N. L. Pedersen (2006). "Role of genes and environments for explaining Alzheimer disease." Arch Gen Psychiatry **63**(2): 168-174.

Gentleman, R. C., V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. Yang and J. Zhang (2004). "Bioconductor: open software development for computational biology and bioinformatics." Genome Biol **5**(10): R80.

George, J. L., S. Mok, D. Moses, S. Wilkins, A. I. Bush, R. A. Cherny and D. I. Finkelstein (2009). "Targeting the Progression of Parkinson's Disease." Current Neuropharmacology **7**(1): 9-36.

Godbout, J. P., J. Chen, J. Abraham, A. F. Richwine, B. M. Berg, K. W. Kelley and R. W. Johnson (2005). "Exaggerated neuroinflammation and sickness behavior in aged mice following activation of the peripheral innate immune system." Faseb j **19**(10): 1329-1331.

Godderis, L., C. Schouteden, A. Tabish, K. Poels, P. Hoet, A. A. Baccarelli and K. Van Landuyt (2015). "Global Methylation and Hydroxymethylation in DNA from Blood and Saliva in Healthy Volunteers." BioMed Research International **2015**: 8.

Golebiewska, A., S. P. Atkinson, M. Lako and L. Armstrong (2009). "Epigenetic landscaping during hESC differentiation to neural cells." Stem Cells **27**(6): 1298-1308.

Gomperts, S. N., D. M. Rentz, E. Moran, J. A. Becker, J. J. Locascio, W. E. Klunk, C. A. Mathis, D. R. Elmaleh, T. Shoup, A. J. Fischman, B. T. Hyman, J. H. Growdon and K. A. Johnson (2008). "Imaging amyloid deposition in Lewy body diseases." Neurology **71**(12): 903-910.

Gorelick, P. B. (2004). "Risk Factors for Vascular Dementia and Alzheimer Disease." Stroke **35**(11 suppl 1): 2620-2622.

Gotz, J., F. Chen, J. van Dorpe and R. M. Nitsch (2001). "Formation of neurofibrillary tangles in P301l tau transgenic mice induced by Abeta 42 fibrils." Science **293**(5534): 1491-1495.

Grant, W. B., A. Campbell, R. F. Itzhaki and J. Savory (2002). "The significance of environmental factors in the etiology of Alzheimer's disease." J Alzheimers Dis **4**(3): 179-189.

Grupe, A., R. Abraham, Y. Li, C. Rowland, P. Hollingworth, A. Morgan, L. Jehu, R. Segurado, D. Stone, E. Schadt, M. Karnoub, P. Nowotny, K. Tacey, J. Catanese, J. Sninsky, C. Brayne, D. Rubinsztein, M. Gill, B. Lawlor, S. Lovestone, P. Holmans, M. O'Donovan, J. C. Morris, L. Thal, A. Goate, M. J. Owen and J. Williams (2007). "Evidence for novel susceptibility genes for late-onset Alzheimer's disease from a genome-wide association study of putative functional variants." Human Molecular Genetics **16**(8): 865-873.

Guan, J.-S., S. J. Haggarty, E. Giacometti, J.-H. Dannenberg, N. Joseph, J. Gao, T. J. F. Nieland, Y. Zhou, X. Wang, R. Mazitschek, J. E. Bradner, R. A. DePinho, R. Jaenisch and L.-H. Tsai (2009). "HDAC2 negatively regulates memory formation and synaptic plasticity." Nature **459**: 55.

Guerreiro, R., A. Wojtas, J. Bras, M. Carrasquillo, E. Rogaeva, E. Majounie, C. Cruchaga, C. Sassi, J. S. Kauwe, S. Younkin, L. Hazrati, J. Collinge, J. Pocock, T. Lashley, J. Williams, J. C. Lambert, P.

Amouyel, A. Goate, R. Rademakers, K. Morgan, J. Powell, P. St George-Hyslop, A. Singleton and J. Hardy (2012). "TREM2 Variants in Alzheimer's Disease." N Engl J Med.

Guerreiro, R., A. Wojtas, J. Bras, M. Carrasquillo, E. Rogaeva, E. Majounie, C. Cruchaga, C. Sassi, J. S. K. Kauwe, S. Younkin, L. Hazrati, J. Collinge, J. Pocock, T. Lashley, J. Williams, J.-C. Lambert, P. Amouyel, A. Goate, R. Rademakers, K. Morgan, J. Powell, P. St. George-Hyslop, A. Singleton and J. Hardy (2013). "TREM2 Variants in Alzheimer's Disease." New England Journal of Medicine **368**(2): 117-127.

Guez-Barber, D., S. Fanous, B. K. Harvey, Y. Zhang, E. Lehrmann, K. G. Becker, M. R. Picciotto and B. T. Hope (2012). "FACS purification of immunolabeled cell types from adult rat brain." J Neurosci Methods **203**(1): 10-18.

Guintivano, J., M. Aryee and Z. Kaminsky (2013). "A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression." Epigenetics **8**(3): 290-302.

Guo, J. U., Y. Su, C. Zhong, G. L. Ming and H. Song (2011). "Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain." Cell **145**(3): 423-434.

Hansen, K. H., A. P. Bracken, D. Pasini, N. Dietrich, S. S. Gehani, A. Monrad, J. Rappsilber, M. Lerdrup and K. Helin (2008). "A model for transmission of the H3K27me3 epigenetic mark." Nature Cell Biology **10**: 1291.

Harald Neumann, M. J. D. (2013). "Variant TREM2 as Risk Factor for Alzheimer's Disease." The New England Journal of Medicine **1**.

Harder, M. N., R. Ribel-Madsen, J. M. Justesen, T. Sparso, E. A. Andersson, N. Grarup, T. Jorgensen, A. Linneberg, T. Hansen and O. Pedersen (2013). "Type 2 diabetes risk alleles near BCAR1 and in ANK1 associate with decreased beta-cell function whereas risk alleles near ANKRD55 and GRB14 associate with decreased insulin sensitivity in the Danish Inter99 cohort." J Clin Endocrinol Metab **98**(4): E801-806.

Hardy, J. and G. Higgins (1992). "Alzheimer's disease: the amyloid cascade hypothesis." Science **256**(5054): 184-185.

Harel, A., F. Wu, M. P. Mattson, C. M. Morris and P. J. Yao (2008). "Evidence for CALM in directing VAMP2 trafficking." Traffic **9**(3): 417-429.

Harikumar, A. and E. Meshorer (2015). "Chromatin remodeling and bivalent histone modifications in embryonic stem cells." EMBO Rep **16**(12): 1609-1619.

Harold, D., R. Abraham, P. Hollingworth, R. Sims, A. Gerrish, M. L. Hamshere, J. S. Pahwa, V. Moskvina, K. Dowzell, A. Williams, N. Jones, C. Thomas, A. Stretton, A. R. Morgan, S. Lovestone, J. Powell, P. Proitsi, M. K. Lupton, C. Brayne, D. C. Rubinsztein, M. Gill, B. Lawlor, A. Lynch, K. Morgan, K. S. Brown, P. A. Passmore, D. Craig, B. McGuinness, S. Todd, C. Holmes, D. Mann, A. D. Smith, S. Love, P. G. Kehoe, J. Hardy, S. Mead, N. Fox, M. Rossor, J. Collinge, W. Maier, F. Jessen, B. Schurmann, R. Heun, H. van den Bussche, I. Heuser, J. Kornhuber, J. Wiltfang, M. Dichgans, L. Frolich, H. Hampel, M. Hull, D. Rujescu, A. M. Goate, J. S. K. Kauwe, C. Cruchaga, P. Nowotny, J. C. Morris, K. Mayo, K. Sleegers, K. Bettens, S. Engelborghs, P. P. De Deyn, C. Van Broeckhoven, G. Livingston, N. J. Bass, H. Gurling, A. McQuillin, R. Gwilliam, P. Deloukas, A. Al-Chalabi, C. E. Shaw, M. Tsolaki, A. B. Singleton, R. Guerreiro, T. W. Muhleisen, M. M. Nothen, S. Moebus, K.-H. Jockel, N. Klopp, H. E. Wichmann, M. M. Carrasquillo, V. S. Pankratz, S. G. Younkin, P. A. Holmans, M. O'Donovan, M. J. Owen and J. Williams (2009). "Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease." Nat Genet **41**(10): 1088-1093.

Harold, D., R. Abraham, P. Hollingworth, R. Sims, A. Gerrish, M. L. Hamshere, J. S. Pahwa, V. Moskvina, K. Dowzell, A. Williams, N. Jones, C. Thomas, A. Stretton, A. R. Morgan, S. Lovestone, J. Powell, P. Proitsi, M. K. Lupton, C. Brayne, D. C. Rubinsztein, M. Gill, B. Lawlor, A. Lynch, K. Morgan, K. S. Brown, P. A. Passmore, D. Craig, B. McGuinness, S. Todd, C. Holmes, D. Mann, A. D. Smith, S. Love, P. G. Kehoe, J. Hardy, S. Mead, N. Fox, M. Rossor, J. Collinge, W. Maier, F. Jessen, B. Schurmann, H. van den Bussche, I. Heuser, J. Kornhuber, J.

Wiltfang, M. Dichgans, L. Frolich, H. Hampel, M. Hull, D. Rujescu, A. M. Goate, J. S. Kauwe, C. Cruchaga, P. Nowotny, J. C. Morris, K. Mayo, K. Sleegers, K. Bettens, S. Engelborghs, P. P. De Deyn, C. Van Broeckhoven, G. Livingston, N. J. Bass, H. Gurling, A. McQuillin, R. Gwilliam, P. Deloukas, A. Al-Chalabi, C. E. Shaw, M. Tsolaki, A. B. Singleton, R. Guerreiro, T. W. Muhleisen, M. M. Nothen, S. Moebus, K. H. Jockel, N. Klopp, H. E. Wichmann, M. M. Carrasquillo, V. S. Pankratz, S. G. Younkin, P. A. Holmans, M. O'Donovan, M. J. Owen and J. Williams (2009). "Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease." Nat Genet **41**(10): 1088-1093.

Harrington, C. T., E. I. Lin, M. T. Olson and J. R. Eshleman (2013). "Fundamentals of pyrosequencing." Arch Pathol Lab Med **137**(9): 1296-1303.

Haruhiko Akiyama, S. B., Scott Barnum, Bonnie Bradt, Joachim Bauer, Greg M. Cole, Neil R. Cooper, Piet Eikelenboom, Mark Emmerling, Berndt L. Fiebich, Caleb E. Finch, Sally Frautschy, W.S.T. Griffin, Harald Hampel, Michael Hull, Gary Landreth, Lih–Fen Lue, Robert Mrak, Ian R. Mackenzie, Patrick L. McGeer, M. Kerry O'Banion, Joel Pachter, Guilio Pasinetti, Carlos Plata–Salaman, Joseph Rogers, Russell Rydel, Yong Shen, Wolfgang Streit, Ronald Strohmeyer, Ikuo Tooyoma, Freek L. Van Muiswinkel, Robert Veerhuis, Douglas Walker, Scott Webster, Beatrice Wegrzyniak, Gary Wenk, Tony Wyss–Coray (2000). "Inflammation and Alzheimer's disease." Neurobiology of Aging **21**: 383–421.

Hashimoto, H., Y. Liu, A. K. Upadhyay, Y. Chang, S. B. Howerton, P. M. Vertino, X. Zhang and X. Cheng (2012). "Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation." Nucleic Acids Research **40**(11): 4841-4849.

Heintzman, N. D., G. C. Hon, R. D. Hawkins, P. Kheradpour, A. Stark, L. F. Harp, Z. Ye, L. K. Lee, R. K. Stuart, C. W. Ching, K. A. Ching, J. E. Antosiewicz, H. Liu, X. Zhang, R. D. Green, R. Stewart, J. A. Thomson, G. E. Crawford, M. Kellis and B. Ren (2009). "Histone Modifications at Human Enhancers Reflect Global Cell Type-Specific Gene Expression." Nature **459**(7243): 108-112.

Hellman, A. and A. Chess (2007). "Gene body-specific methylation on the active X chromosome." Science **315**(5815): 1141-1143.

Hernandez, H. G., M. F. Mahecha, A. Mejia, H. Arboleda and D. A. Forero (2014). "Global long interspersed nuclear element 1 DNA methylation in a Colombian sample of patients with late-onset Alzheimer's disease." Am J Alzheimers Dis Other Demen **29**(1): 50-53.

Hollingworth, P., D. Harold, R. Sims, A. Gerrish, J. C. Lambert, M. M. Carrasquillo, R. Abraham, M. L. Hamshere, J. S. Pahwa, V. Moskvina, K. Dowzell, N. Jones, A. Stretton, C. Thomas, A. Richards, D. Ivanov, C. Widdowson, J. Chapman, S. Lovestone, J. Powell, P. Proitsi, M. K. Lupton, C. Brayne, D. C. Rubinsztein, M. Gill, B. Lawlor, A. Lynch, K. S. Brown, P. A. Passmore, D. Craig, B. McGuinness, S. Todd, C. Holmes, D. Mann, A. D. Smith, H. Beaumont, D. Warden, G. Wilcock, S. Love, P. G. Kehoe, N. M. Hooper, E. R. Vardy, J. Hardy, S. Mead, N. C. Fox, M. Rossor, J. Collinge, W. Maier, F. Jessen, E. Ruther, B. Schurmann, R. Heun, H. Kolsch, H. van den Bussche, I. Heuser, J. Kornhuber, J. Wiltfang, M. Dichgans, L. Frolich, H. Hampel, J. Gallacher, M. Hull, D. Rujescu, I. Giegling, A. M. Goate, J. S. Kauwe, C. Cruchaga, P. Nowotny, J. C. Morris, K. Mayo, K. Sleegers, K. Bettens, S. Engelborghs, P. P. De Deyn, C. Van Broeckhoven, G. Livingston, N. J. Bass, H. Gurling, A. McQuillin, R. Gwilliam, P. Deloukas, A. Al-Chalabi, C. E. Shaw, M. Tsolaki, A. B. Singleton, R. Guerreiro, T. W. Muhleisen, M. M. Nothen, S. Moebus, K. H. Jockel, N. Klopp, H. E. Wichmann, V. S. Pankratz, S. B. Sando, J. O. Aasly, M. Barcikowska, Z. K. Wszolek, D. W. Dickson, N. R. Graff-Radford, R. C. Petersen, C. M. van Duijn, M. M. Breteler, M. A. Ikram, A. L. DeStefano, A. L. Fitzpatrick, O. Lopez, L. J. Launer, S. Seshadri, C. Berr, D. Campion, J. Epelbaum, J. F. Dartigues, C. Tzourio, A. Alperovitch, M. Lathrop, T. M. Feulner, P. Friedrich, C. Riehle, M. Krawczak, S. Schreiber, M. Mayhaus, S. Nicolhaus, S. Wagenpfeil, S. Steinberg, H. Stefansson, K. Stefansson, J. Snaedal, S. Bjornsson, P. V. Jonsson, V. Chouraki, B. Genier-Boley, M. Hiltunen, H. Soininen, O. Combarros, D. Zelenika, M. Delepine, M. J. Bullido, F. Pasquier, I. Mateo, A. Frank-Garcia, E. Porcellini, O. Hanon, E. Coto, V. Alvarez, P. Bosco, G. Siciliano, M. Mancuso, F. Panza, V.

Solfrizzi, B. Nacmias, S. Sorbi, P. Bossu, P. Piccardi, B. Arosio, G. Annoni, D. Seripa, A. Pilotto, E. Scarpini, D. Galimberti, A. Brice, D. Hannequin, F. Licastro, L. Jones, P. A. Holmans, T. Jonsson, M. Riemenschneider, K. Morgan, S. G. Younkin, M. J. Owen, M. O'Donovan, P. Amouyel and J. Williams (2011). "Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease." Nat Genet **43**(5): 429-435.

Hu, X., E. Pickering, Y. C. Liu, S. Hall, H. Fournier, E. Katz, B. Dechairo, S. John, P. Van Eerdewegh and H. Soares (2011). "Meta-analysis for genome-wide association study identifies multiple variants at the BIN1 locus associated with late-onset Alzheimer's disease." PLoS One **6**(2): e16616.

Huang, Y., W. A. Pastor, Y. Shen, M. Tahiliani, D. R. Liu and A. Rao (2010). "The Behaviour of 5-Hydroxymethylcytosine in Bisulfite Sequencing." PLOS ONE **5**(1): e8888.

Hughes, M. R., N. Anderson, S. Maltby, J. Wong, Z. Berberovic, C. S. Birkenmeier, D. J. Haddon, K. Garcha, A. Flenniken, L. R. Osborne, S. L. Adamson, J. Rossant, L. L. Peters, M. D. Minden, R. F. Paulson, C. Wang, D. L. Barber, K. M. McNagny and W. L. Stanford (2011). "A novel ENU-generated truncation mutation lacking the spectrin-binding and C-terminal regulatory domains of Ank1 models severe hemolytic hereditary spherocytosis." Experimental Hematology **39**(3): 305-320.e302.

Hunter, C. A., N. Y. Kirson, U. Desai, A. K. Cummings, D. E. Faries and H. G. Birnbaum (2015). "Medical costs of Alzheimer's disease misdiagnosis among US Medicare beneficiaries." Alzheimers Dement **11**(8): 887-895.

Hwang, I. K., C. H. Lee, H. Li, K. Y. Yoo, J. H. Choi, D. W. Kim, D. W. Kim, H. W. Suh and M. H. Won (2008). "Comparison of ionized calcium-binding adapter molecule 1 immunoreactivity of the hippocampal dentate gyrus and CA1 region in adult and aged dogs." Neurochem Res **33**(7): 1309-1315.

Illumina. (2015). "Infinium HD Assay Methylation Protocol Guide." Retrieved 16.08.2017, 2017.

Imagilys (2015). Neuroimaging Biomarkers for ALzheimer's Disease. Online.

Imamura, M., S. Maeda, T. Yamauchi, K. Hara, K. Yasuda, T. Morizono, A. Takahashi, M. Horikoshi, M. Nakamura, H. Fujita, T. Tsunoda, M. Kubo, H. Watada, H. Maegawa, M. Okada-Iwabu, M. Iwabu, N. Shojima, T. Ohshige, S. Omori, M. Iwata, H. Hirose, K. Kaku, C. Ito, Y. Tanaka, K. Tobe, A. Kashiwagi, R. Kawamori, M. Kasuga, N. Kamatani, Y. Nakamura and T. Kadowaki (2012). "A single-nucleotide polymorphism in ANK1 is associated with susceptibility to type 2 diabetes in Japanese populations." Hum Mol Genet **21**(13): 3042-3049.

Inoue, A., L. Shen, Q. Dai, C. He and Y. Zhang (2011). "Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development." Cell Res **21**(12): 1670-1676.

Inoue, A. and Y. Zhang (2011). "Replication-Dependent Loss of 5-Hydroxymethylcytosine in Mouse Preimplantation Embryos." Science **334**(6053): 194-194.

Ito, S., A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers and Y. Zhang (2010). "Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification." Nature **466**.

Ito, S., L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He and Y. Zhang (2011). "Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine." Science **333**(6047): 1300-1303.

Iurlaro, M., G. Ficz, D. Oxley, E.-A. Raiber, M. Bachman, M. J. Booth, S. Andrews, S. Balasubramanian and W. Reik (2013). "A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation." Genome Biology **14**(10): R119.

Iwata, A., K. Nagata, H. Hatsuta, H. Takuma, M. Bundo, K. Iwamoto, A. Tamaoka, S. Murayama, T. Saido and S. Tsuji (2014). "Altered CpG methylation in sporadic Alzheimer's disease is associated with APP and MAPT dysregulation." Hum Mol Genet **23**(3): 648-656.

Jack, C. R., D. S. Knopman, W. J. Jagust, L. M. Shaw, P. S. Aisen, M. W. Weiner, R. C. Petersen and J. Q. Trojanowski (2010). "Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade." Lancet neurology **9**(1): 119.

Jacobsen, L., P. Madsen, C. Jacobsen, M. S. Nielsen, J. Gliemann and C. M. Petersen (2001). "Activation and functional characterization of the mosaic receptor SorLA/LR11." J Biol Chem **276**(25): 22788-22796.

Jaenisch, R. and A. Bird (2003). "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals." Nat Genet **33 Suppl**: 245-254.

Jankovic, J. (2008). "Parkinson's disease: clinical features and diagnosis." Journal of Neurology, Neurosurgery & Psychiatry **79**(4): 368-376.

Janson, J., T. Laedtke, J. E. Parisi, P. O'Brien, R. C. Petersen and P. C. Butler (2004). "Increased Risk of Type 2 Diabetes in Alzheimer Disease." Diabetes **53**(2): 474-481.

Jin, B., Y. Li and K. D. Robertson (2011). "DNA methylation: superior or subordinate in the epigenetic hierarchy?" Genes Cancer **2**(6): 607-617.

Jones, P. A. (2012). "Functions of DNA methylation: islands, start sites, gene bodies and beyond." Nat Rev Genet **13**.

Jones, P. A. and P. W. Laird (1999). "Cancer epigenetics comes of age." Nat Genet **21**(2): 163-167.

Jonsson, T., H. Stefansson, S. Steinberg, I. Jonsdottir, P. V. Jonsson, J. Snaedal, S. Bjornsson, J. Huttenlocher, A. I. Levey, J. J. Lah, D. Rujescu, H. Hampel, I. Giegling, O. A. Andreassen, K. Engedal, I. Ulstein, S. Djurovic, C. Ibrahim-Verbaas, A. Hofman, M. A. Ikram, C. M. van Duijn, U. Thorsteinsdottir, A. Kong and K. Stefansson (2013). "Variant of TREM2 Associated with the Risk of Alzheimer's Disease." New England Journal of Medicine **368**(2): 107-116.

Joutel, A., C. Corpechot, A. Ducros, K. Vahedi, H. Chabriat, P. Mouton, S. Alamowitch, V. Domenga, M. Cecillion, E. Marechal, J. Maciazek, C. Vayssiere, C. Cruaud, E. A. Cabanis, M. M. Ruchoux, J. Weissenbach, J. F. Bach, M. G. Bousser and E. Tournier-Lasserve (1996). "Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia." Nature **383**(6602): 707-710.

Jun, G., A. C. Naj, G. W. Beecham, L. S. Wang, J. Buros, P. J. Gallins, J. D. Buxbaum, N. Ertekin-Taner, M. D. Fallin, R. Friedland, R. Inzelberg, P. Kramer, E. Rogaeva, P. St George-Hyslop, L. B. Cantwell, B. A. Dombroski, A. J. Saykin, E. M. Reiman, D. A. Bennett, J. C. Morris, K. L. Lunetta, E. R. Martin, T. J. Montine, A. M. Goate, D. Blacker, D. W. Tsuang, D. Beekly, L. A. Cupples, H. Hakonarson, W. Kukull, T. M. Foroud, J. Haines, R. Mayeux, L. A. Farrer, M. A. Pericak-Vance and G. D. Schellenberg (2010). "Meta-analysis confirms CR1, CLU, and PICALM as alzheimer disease risk loci and reveals interactions with APOE genotypes." Arch Neurol **67**(12): 1473-1484.

Kalaria, R. N. and C. Ballard (1999). "Overlap between pathology of Alzheimer disease and vascular dementia." Alzheimer Dis Assoc Disord **13 Suppl 3**: S115-123.

Kamboh, M. I., F. Y. Demirci, X. Wang, R. L. Minster, M. M. Carrasquillo, V. S. Pankratz, S. G. Younkin, A. J. Saykin, I. Alzheimer's Disease Neuroimaging, G. Jun, C. Baldwin, M. W. Logue, J. Buros, L. Farrer, M. A. Pericak-Vance, J. L. Haines, R. A. Sweet, M. Ganguli, E. Feingold, S. T. Dekosky, O. L. Lopez and M. M. Barmada (2012). "Genome-wide association study of Alzheimer's disease." Transl Psychiatry **2**: e117.

Kapushesky, M., I. Emam, E. Holloway, P. Kurnosov, A. Zorin, J. Malone, G. Rustici, E. Williams, H. Parkinson and A. Brazma (2010). "Gene Expression Atlas at the European Bioinformatics Institute." Nucleic Acids Research **38**(Database issue): D690-D698.

Karlić, R., H.-R. Chung, J. Lasserre, K. Vlahoviček and M. Vingron (2010). "Histone modification levels are predictive for gene expression." Proceedings of the National Academy of Sciences **107**(7): 2926-2931.

Kayne, P. S., U. J. Kim, M. Han, J. R. Mullen, F. Yoshizaki and M. Grunstein (1988). "Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast." Cell **55**(1): 27-39.

Khare, T., S. Pai, K. Koncevicius, M. Pal, E. Kriukiene, Z. Liutkeviciute, M. Irimia, P. Jia, C. Ptak, M. Xia, R. Tice, M. Tochigi, S. Morera, A. Nazarians, D. Belsham, A. H. Wong, B. J. Blencowe, S. C. Wang, P. Kapranov, R. Kustra, V. Labrie, S. Klimasauskas and A. Petronis (2012). "5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary." Nat Struct Mol Biol **19**(10): 1037-1043.

Khatri, P., M. Sirota and A. J. Butte (2012). "Ten Years of Pathway Analysis: Current Approaches and Outstanding Challenges." PLoS Computational Biology **8**(2): e1002375.

Kim, H. S., E. M. Kim, N. J. Kim, K. A. Chang, Y. Choi, K. W. Ahn, J. H. Lee, S. Kim, C. H. Park and Y. H. Suh (2004). "Inhibition of histone deacetylation enhances the neurotoxicity induced by the c-terminal fragments of amyloid precursor protein." Journal of neuroscience research **75**(1): 117-124.

Kim, Y. Z. (2014). "Altered Histone Modifications in Gliomas." Brain Tumor Res Treat **2**(1): 7-21.

Klose, R. J. and A. P. Bird (2006). "Genomic DNA methylation: the mark and its mediators." Trends Biochem Sci **31**(2): 89-97.

Kondo, Y. (2009). "Epigenetic Cross-Talk between DNA Methylation and Histone Modifications in Human Cancers." Yonsei Medical Journal **50**(4): 455-463.

Kosik, K. S., C. L. Joachim and D. J. Selkoe (1986). "Microtubule-associated protein tau (tau) is a major antigenic component of paired helical filaments in Alzheimer disease." Proc Natl Acad Sci U S A **83**(11): 4044-4048.

Kouzarides, T. (2007). "Chromatin modifications and their function." Cell **128**.

Kozomara, A. and S. Griffiths-Jones (2014). "miRBase: annotating high confidence microRNAs using deep sequencing data." Nucleic Acids Research **42**(D1): D68-D73.

Kramer, P. L., H. Xu, R. L. Woltjer, S. K. Westaway, D. Clark, D. Erten-Lyons, J. A. Kaye, K. A. Welsh-Bohmer, J. C. Troncoso, W. R. Markesbery, R. C. Petersen, R. S. Turner, W. A. Kukull, D. A. Bennett, D. Galasko, J. C. Morris and J. Ott (2011). "Alzheimer disease pathology in cognitively healthy elderly: a genome-wide study." Neurobiol Aging **32**(12): 2113-2122.

Kriaucionis, S. and N. Heintz (2009). "The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain." Science **324**(5929): 929-930.

Kriaucionis, S. and N. Heintz (2009). "The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain." Science **324**(5929): 929-930.

Krishnaswami, S. R., R. V. Grindberg, M. Novotny, P. Venepally, B. Lacar, K. Bhutani, S. B. Linker, S. Pham, J. A. Erwin, J. A. Miller, R. Hodge, J. K. McCarthy, M. Kelder, J. McCorrison, B. D. Aevermann, F. D. Fuertes, R. H. Scheuermann, J. Lee, E. S. Lein, N. Schork, M. J. McConnell, F. H. Gage and R. S. Lasken (2016). "Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons." Nature protocols **11**(3): 499-524.

Kuo, Y. M., M. R. Emmerling, C. L. Bisgaier, A. D. Essenburg, H. C. Lampert, D. Drumm and A. E. Roher (1998). "Elevated low-density lipoprotein in Alzheimer's disease correlates with brain abeta 1-42 levels." Biochem Biophys Res Commun **252**(3): 711-715.

LaDu, M. J., M. T. Falduto, A. M. Manelli, C. A. Reardon, G. S. Getz and D. E. Frail (1994). "Isoform-specific binding of apolipoprotein E to beta-amyloid." Journal of Biological Chemistry **269**(38): 23403-23406.

Laird, P. W. (2005). "Cancer epigenetics." Human Molecular Genetics **14**(suppl 1): R65-R76.

Lambert, J.-C., C. A. Ibrahim-Verbaas, D. Harold, A. C. Naj, R. Sims, C. Bellenguez, G. Jun, A. L. DeStefano, J. C. Bis, G. W. Beecham, B. Grenier-Boley, G. Russo, T. A. Thornton-Wells, N. Jones, A. V. Smith, V. Chouraki, C. Thomas, M. A. Ikram, D. Zelenika, B. N. Vardarajan, Y. Kamatani, C.-F. Lin, A. Gerrish, H. Schmidt, B. Kunkle, M. L. Dunstan, A. Ruiz, M.-T. Bihoreau, S.-H. Choi, C. Reitz, F. Pasquier, P. Hollingworth, A. Ramirez, O. Hanon, A. L. Fitzpatrick, J. D. Buxbaum, D. Campion, P. K. Crane, C. Baldwin, T. Becker, V. Gudnason, C. Cruchaga, D. Craig, N. Amin, C. Berr, O. L. Lopez, P. L. De Jager, V. Deramecourt, J. A. Johnston, D. Evans, S. Lovestone, L. Letenneur, F. J. Moron, D. C. Rubinsztein, G. Eiriksdottir, K. Sleegers, A. M. Goate, N. Fievet, M. J. Huentelman, M. Gill, K. Brown, M. I. Kamboh, L. Keller, P. Barberger-

Gateau, B. McGuinness, E. B. Larson, R. Green, A. J. Myers, C. Dufouil, S. Todd, D. Wallon, S. Love, E. Rogaeva, J. Gallacher, P. St George-Hyslop, J. Clarimon, A. Lleo, A. Bayer, D. W. Tsuang, L. Yu, M. Tsolaki, P. Bossu, G. Spalletta, P. Proitsi, J. Collinge, S. Sorbi, F. Sanchez-Garcia, N. C. Fox, J. Hardy, M. C. D. Naranjo, P. Bosco, R. Clarke, C. Brayne, D. Galimberti, M. Mancuso, F. Matthews, I. European Alzheimer's Disease, Genetic, D. Environmental Risk in Alzheimer's, C. Alzheimer's Disease Genetic, H. Cohorts for, E. Aging Research in Genomic, S. Moebus, P. Mecocci, M. Del Zompo, W. Maier, H. Hampel, A. Pilotto, M. Bullido, F. Panza, P. Caffarra, B. Nacmias, J. R. Gilbert, M. Mayhaus, L. Lannfelt, H. Hakonarson, S. Pichler, M. M. Carrasquillo, M. Ingelsson, D. Beekly, V. Alvarez, F. Zou, O. Valladares, S. G. Younkin, E. Coto, K. L. Hamilton-Nelson, W. Gu, C. Razquin, P. Pastor, I. Mateo, M. J. Owen, K. M. Faber, P. V. Jonsson, O. Combarros, M. C. O'Donovan, L. B. Cantwell, H. Soininen, D. Blacker, S. Mead, T. H. Mosley Jr, D. A. Bennett, T. B. Harris, L. Fratiglioni, C. Holmes, R. F. A. G. de Bruijn, P. Passmore, T. J. Montine, K. Bettens, J. I. Rotter, A. Brice, K. Morgan, T. M. Foroud, W. A. Kukull, D. Hannequin, J. F. Powell, M. A. Nalls, K. Ritchie, K. L. Lunetta, J. S. K. Kauwe, E. Boerwinkle, M. Riemenschneider, M. Boada, M. Hiltunen, E. R. Martin, R. Schmidt, D. Rujescu, L.-S. Wang, J.-F. Dartigues, R. Mayeux, C. Tzourio, A. Hofman, M. M. Nothen, C. Graff, B. M. Psaty, L. Jones, J. L. Haines, P. A. Holmans, M. Lathrop, M. A. Pericak-Vance, L. J. Launer, L. A. Farrer, C. M. van Duijn, C. Van Broeckhoven, V. Moskvina, S. Seshadri, J. Williams, G. D. Schellenberg and P. Amouyel (2013). "Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease." Nat Genet **45**(12): 1452-1458.

Lambert, J. C., S. Heath, G. Even, D. Campion, K. Sleegers, M. Hiltunen, O. Combarros, D. Zelenika, M. J. Bullido, B. Tavernier, L. Letenneur, K. Bettens, C. Berr, F. Pasquier, N. Fievet, P. Barberger-Gateau, S. Engelborghs, P. De Deyn, I. Mateo, A. Franck, S. Helisalmi, E. Porcellini, O. Hanon, M. M. de Pancorbo, C. Lendon, C. Dufouil, C. Jaillard, T. Leveillard, V. Alvarez, P. Bosco, M. Mancuso, F. Panza, B. Nacmias, P. Bossu, P. Piccardi, G. Annoni, D. Seripa, D. Galimberti, D. Hannequin, F. Licastro, H. Soininen, K. Ritchie, H. Blanche, J. F. Dartigues, C. Tzourio, I. Gut, C. Van Broeckhoven, A. Alperovitch, M. Lathrop and P. Amouyel (2009). "Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease." Nat Genet **41**(10): 1094-1099.

Lambert, J. C., C. A. Ibrahim-Verbaas, D. Harold, A. C. Naj, R. Sims, C. Bellenguez, A. L. DeStafano, J. C. Bis, G. W. Beecham, B. Grenier-Boley, G. Russo, T. A. Thorton-Wells, N. Jones, A. V. Smith, V. Chouraki, C. Thomas, M. A. Ikram, D. Zelenika, B. N. Vardarajan, Y. Kamatani, C. F. Lin, A. Gerrish, H. Schmidt, B. Kunkle, M. L. Dunstan, A. Ruiz, M. T. Bihoreau, S. H. Choi, C. Reitz, F. Pasquier, C. Cruchaga, D. Craig, N. Amin, C. Berr, O. L. Lopez, P. L. De Jager, V. Deramecourt, J. A. Johnston, D. Evans, S. Lovestone, L. Letenneur, F. J. Moron, D. C. Rubinsztein, G. Eiriksdottir, K. Sleegers, A. M. Goate, N. Fievet, M. W. Huentelman, M. Gill, K. Brown, M. I. Kamboh, L. Keller, P. Barberger-Gateau, B. McGuiness, E. B. Larson, R. Green, A. J. Myers, C. Dufouil, S. Todd, D. Wallon, S. Love, E. Rogaeva, J. Gallacher, P. St George-Hyslop, J. Clarimon, A. Lleo, A. Bayer, D. W. Tsuang, L. Yu, M. Tsolaki, P. Bossu, G. Spalletta, P. Proitsi, J. Collinge, S. Sorbi, F. Sanchez-Garcia, N. C. Fox, J. Hardy, M. C. Deniz Naranjo, P. Bosco, R. Clarke, C. Brayne, D. Galimberti, M. Mancuso, F. Matthews, I. European Alzheimer's Disease, Genetic, D. Environmental Risk in Alzheimer's, C. Alzheimer's Disease Genetic, H. Cohorts for, E. Aging Research in Genomic, S. Moebus, P. Mecocci, M. Del Zompo, W. Maier, H. Hampel, A. Pilotto, M. Bullido, F. Panza, P. Caffarra, B. Nacmias, J. R. Gilbert, M. Mayhaus, L. Lannefelt, H. Hakonarson, S. Pichler, M. M. Carrasquillo, M. Ingelsson, D. Beekly, V. Alvarez, F. Zou, O. Valladares, S. G. Younkin, E. Coto, K. L. Hamilton-Nelson, W. Gu, C. Razquin, P. Pastor, I. Mateo, M. J. Owen, K. M. Faber, P. V. Jonsson, O. Combarros, M. C. O'Donovan, L. B. Cantwell, H. Soininen, D. Blacker, S. Mead, T. H. Mosley, Jr., D. A. Bennett, T. B. Harris, L. Fratiglioni, C. Holmes, R. F. de Bruijn, P. Passmore, T. J. Montine, K. Bettens, J. I. Rotter, A. Brice, K. Morgan, T. M. Foroud, W. A. Kukull, D. Hannequin, J. F. Powell, M. A. Nalls, K. Ritchie, K. L. Lunetta, J. S. Kauwe, E. Boerwinkle, M. Riemenschneider, M. Boada, M.

Hiltuenen, E. R. Martin, R. Schmidt, D. Rujescu, L. S. Wang, J. F. Dartigues, R. Mayeux, C. Tzourio, A. Hofman, M. M. Nothen, C. Graff, B. M. Psaty, L. Jones, J. L. Haines, P. A. Holmans, M. Lathrop, M. A. Pericak-Vance, L. J. Launer, L. A. Farrer, C. M. van Duijn, C. Van Broeckhoven, V. Moskvina, S. Seshadri, J. Williams, G. D. Schellenberg and P. Amouyel (2013). "Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease." Nat Genet **45**(12): 1452-1458.

Lambert, J. C., C. A. Ibrahim-Verbaas, D. Harold, A. C. Naj, R. Sims, C. Bellenguez, G. Jun, A. L. Destefano, J. C. Bis, G. W. Beecham, B. Grenier-Boley, G. Russo, T. A. Thornton-Wells, N. Jones, A. V. Smith, V. Chouraki, C. Thomas, M. A. Ikram, D. Zelenika, B. N. Vardarajan, Y. Kamatani, C. F. Lin, A. Gerrish, H. Schmidt, B. Kunkle, M. L. Dunstan, A. Ruiz, M. T. Bihoreau, S. H. Choi, C. Reitz, F. Pasquier, P. Hollingworth, A. Ramirez, O. Hanon, A. L. Fitzpatrick, J. D. Buxbaum, D. Campion, P. K. Crane, C. Baldwin, T. Becker, V. Gudnason, C. Cruchaga, D. Craig, N. Amin, C. Berr, O. L. Lopez, P. L. De Jager, V. Deramecourt, J. A. Johnston, D. Evans, S. Lovestone, L. Letenneur, F. J. Moron, D. C. Rubinsztein, G. Eiriksdottir, K. Sleegers, A. M. Goate, N. Fievet, M. J. Huentelman, M. Gill, K. Brown, M. I. Kamboh, L. Keller, P. Barberger-Gateau, B. McGuinness, E. B. Larson, R. Green, A. J. Myers, C. Dufouil, S. Todd, D. Wallon, S. Love, E. Rogaeva, J. Gallacher, P. St George-Hyslop, J. Clarimon, A. Lleo, A. Bayer, D. W. Tsuang, L. Yu, M. Tsolaki, P. Bossu, G. Spalletta, P. Proitsi, J. Collinge, S. Sorbi, F. Sanchez-Garcia, N. C. Fox, J. Hardy, M. C. Naranjo, P. Bosco, R. Clarke, C. Brayne, D. Galimberti, M. Mancuso, F. Matthews, I. European Alzheimer's Disease, Genetic, D. Environmental Risk in Alzheimer's, C. Alzheimer's Disease Genetic, H. Cohorts for, E. Aging Research in Genomic, S. Moebus, P. Mecocci, M. Del Zompo, W. Maier, H. Hampel, A. Pilotto, M. Bullido, F. Panza, P. Caffarra, B. Nacmias, J. R. Gilbert, M. Mayhaus, L. Lannfelt, H. Hakonarson, S. Pichler, M. M. Carrasquillo, M. Ingelsson, D. Beekly, V. Alvarez, F. Zou, O. Valladares, S. G. Younkin, E. Coto, K. L. Hamilton-Nelson, W. Gu, C. Razquin, P. Pastor, I. Mateo, M. J. Owen, K. M. Faber, P. V. Jonsson, O. Combarros, M. C. O'Donovan, L. B. Cantwell, H. Soininen, D. Blacker, S. Mead, T. H. Mosley, Jr., D. A. Bennett, T. B. Harris, L. Fratiglioni, C. Holmes, R. F. de Bruijn, P. Passmore, T. J. Montine, K. Bettens, J. I. Rotter, A. Brice, K. Morgan, T. M. Foroud, W. A. Kukull, D. Hannequin, J. F. Powell, M. A. Nalls, K. Ritchie, K. L. Lunetta, J. S. Kauwe, E. Boerwinkle, M. Riemenschneider, M. Boada, M. Hiltunen, E. R. Martin, R. Schmidt, D. Rujescu, L. S. Wang, J. F. Dartigues, R. Mayeux, C. Tzourio, A. Hofman, M. M. Nothen, C. Graff, B. M. Psaty, L. Jones, J. L. Haines, P. A. Holmans, M. Lathrop, M. A. Pericak-Vance, L. J. Launer, L. A. Farrer, C. M. van Duijn, C. Van Broeckhoven, V. Moskvina, S. Seshadri, J. Williams, G. D. Schellenberg and P. Amouyel (2013). "Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease." Nat Genet **45**(12): 1452-1458.

Lardenoije, R., D. L. van den Hove, T. S. Vaessen, A. Iatrou, K. P. Meuwissen, B. T. van Hagen, G. Kenis, H. W. Steinbusch, C. Schmitz and B. P. Rutten (2015). "Epigenetic modifications in mouse cerebellar Purkinje cells: effects of aging, caloric restriction, and overexpression of superoxide dismutase 1 on 5-methylcytosine and 5-hydroxymethylcytosine." Neurobiol Aging **36**(11): 3079-3089.

Lashley, T., P. Gami, N. Valizadeh, A. Li, T. Revesz and R. Balazs (2015). "Alterations in global DNA methylation and hydroxymethylation are not detected in Alzheimer's disease." Neuropathol Appl Neurobiol **41**(4): 497-506.

Laurent, L., E. Wong, G. Li, T. Huynh, A. Tsirigos, C. T. Ong, H. M. Low, K. W. Kin Sung, I. Rigoutsos, J. Loring and C. L. Wei (2010). "Dynamic changes in the human methylome during differentiation." Genome Res **20**(3): 320-331.

LBDA, L. B. D. A. (2016). "DLB Diagnosis." Retrieved 22 Jan, 2018.

Lee, T. I., R. G. Jenner, L. A. Boyer, M. G. Guenther, S. S. Levine, R. M. Kumar, B. Chevalier, S. E. Johnstone, M. F. Cole, K.-i. Isono, H. Koseki, T. Fuchikami, K. Abe, H. L. Murray, J. P. Zucker, B. Yuan, G. W. Bell, E. Herbolsheimer, N. M. Hannett, K. Sun, D. T. Odom, A. P. Otte, T. L. Volkert, D. P. Bartel, D. A. Melton, D. K. Gifford, R. Jaenisch and R. A. Young (2006). "Control

of Developmental Regulators by Polycomb in Human Embryonic Stem Cells." Cell **125**(2): 301-313.
Leonard, W. (2011). Healthline
Lesch, B. J. and D. C. Page (2014). "Poised chromatin in the mammalian germ line." Development (Cambridge, England) **141**(19): 3619-3626.
Lewis, F., S. Karlsberg Schaffer, J. Sussex, P. O'Neill and L. Cockcroft (2014). The Trajectory of Dementia in the UK - Making a Difference.
Li, L. and C. Hölscher (2007). "Common pathological processes in Alzheimer disease and type 2 diabetes: A review." Brain Research Reviews **56**(2): 384-402.
Li, W. and M. Liu (2011). "Distribution of 5-Hydroxymethylcytosine in Different Human Tissues." Journal of Nucleic Acids **2011**.
Lindsay, J., D. Laurin, R. Verreault, R. Hébert, B. Helliwell, G. B. Hill and I. McDowell (2002). "Risk Factors for Alzheimer's Disease: A Prospective Analysis from the Canadian Study of Health and Aging." American Journal of Epidemiology **156**(5): 445-453.
Lister, R., E. A. Mukamel, J. R. Nery, M. Urich, C. A. Puddifoot and N. D. Johnson (2013). "Global epigenomic reconfiguration during mammalian brain development." Science **341**.
Lister, R., M. Pelizzola, R. H. Dowen, R. D. Hawkins, G. Hon, J. Tonti-Filippini, J. R. Nery, L. Lee, Z. Ye, Q. M. Ngo, L. Edsall, J. Antosiewicz-Bourget, R. Stewart, V. Ruotti, A. H. Millar, J. A. Thomson, B. Ren and J. R. Ecker (2009). "Human DNA methylomes at base resolution show widespread epigenomic differences." Nature **462**(7271): 315-322.
Logue, M. W., M. Schu, B. N. Vardarajan, J. Buros, R. C. Green, R. C. Go, P. Griffith, T. O. Obisesan, R. Shatz, A. Borenstein, L. A. Cupples, K. L. Lunetta, M. D. Fallin, C. T. Baldwin and L. A. Farrer (2011). "A comprehensive genetic association study of Alzheimer disease in African Americans." Arch Neurol **68**(12): 1569-1579.
Louwersheimer, E., A. Ramirez, C. Cruchaga, T. Becker, J. Kornhuber, O. Peters, S. Heilmann, J. Wiltfang, F. Jessen, P. J. Visser, P. Scheltens, Y. A. L. Pijnenburg, C. E. Teunissen, F. Barkhof, J. C. van Swieten, H. Holstege and W. M. Van der Flier (2015). "The influence of genetic variants in SORL1 gene on the manifestation of Alzheimer's disease." Neurobiology of Aging **36**(3): 1605.e1613-1605.e1620.
Lovkvist, C., I. B. Dodd, K. Sneppen and J. O. Haerter (2016). "DNA methylation in human epigenomes depends on local topology of CpG sites." Nucleic Acids Res **44**(11): 5123-5132.
Lu, F., H. Guan, B. Gong, X. Liu, R. Zhu, Y. Wang, J. Qian, T. Zhou, X. Lan, P. Wang, Y. Lin, S. Ma, H. Lin, X. Zhu, R. Chen, X. Zhu, Y. Shi and Z. Yang (2014). "Genetic variants in PVRL2-TOMM40-APOE region are associated with human longevity in a Han Chinese population." PLoS One **9**(6): e99580.
Luchsinger, J. A., M.-X. Tang, Y. Stern, S. Shea and R. Mayeux (2001). "Diabetes Mellitus and Risk of Alzheimer's Disease and Dementia with Stroke in a Multiethnic Cohort." American Journal of Epidemiology **154**(7): 635-641.
Lue, L.-F., Y.-M. Kuo, A. E. Roher, L. Brachova, Y. Shen, L. Sue, T. Beach, J. H. Kurth, R. E. Rydel and J. Rogers (1999). "Soluble Amyloid β Peptide Concentration as a Predictor of Synaptic Change in Alzheimer's Disease." The American Journal of Pathology **155**(3): 853-862.
Lukiw, W. J. (2013). "Alzheimer's disease (AD) as a disorder of the plasma membrane." Frontiers in Physiology **4**: 24.
Lunnon, K., E. Hannon, R. G. Smith, E. Dempster, C. Wong, J. Burrage, C. Troakes, S. Al-Sarraj, A. Kepa, L. Schalkwyk and J. Mill (2016). "Variation in 5-hydroxymethylcytosine across human cortex and cerebellum." Genome Biol **17**: 27.
Lunnon, K., Z. Ibrahim, P. Proitsi, A. Lourdusamy, S. Newhouse, M. Sattlecker, S. Furney, M. Saleem, H. Soininen and I. Kłoszewska (2012). "Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood." Journal of Alzheimer's Disease **30**(3): 685.

Lunnon, K. and J. Mill (2013). "Epigenetic studies in Alzheimer's disease: current findings, caveats, and considerations for future studies." Am J Med Genet B Neuropsychiatr Genet **162B**(8): 789-799.

Lunnon, K., R. Smith, E. Hannon, P. L. De Jager, G. Srivastava, M. Volta, C. Troakes, S. Al-Sarraj, J. Burrage, R. Macdonald, D. Condliffe, L. W. Harries, P. Katsel, V. Haroutunian, Z. Kaminsky, C. Joachim, J. Powell, S. Lovestone, D. A. Bennett, L. C. Schalkwyk and J. Mill (2014). "Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease." Nat Neurosci **17**(9): 1164-1170.

Lunnon, K., R. Smith, E. J. Hannon, P. L. De Jager, G. Srivastava, M. Volta, C. Troakes, S. Al-Sarraj, J. Burrage, R. Macdonald, D. Condliffe, P. Katsel, V. Haroutunian, Z. Kaminsky, C. Joachim, J. Powell, S. Lovestone, D. A. Bennett, L. C. Schalkwyk and J. Mill (2014). "Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease." Nat Neurosci **Sept; 17**(9): 1164-1170.

Lutz, M. W., D. G. Crenshaw, A. M. Saunders and A. D. Roses (2010). "Genetic variation at a single locus and age of onset for Alzheimer's Disease." Alzheimer's & dementia : the journal of the Alzheimer's Association **6**(2): 125-131.

Lyko, F., S. Foret, R. Kucharski, S. Wolf, C. Falckenhayn and R. Maleszka (2010). "The honey bee epigenomes: differential methylation of brain DNA in queens and workers." PLoS Biol **8**(11): e1000506.

Mackenzie, I. R. A. (2000). "Activated microglia in dementia with Lewy bodies." Neurology **55**(1): 132-134.

Marambaud, P., J. Shioi, G. Serban, A. Georgakopoulos, S. Sarner, V. Nagy, L. Baki, P. Wen, S. Efthimiopoulos, Z. Shao, T. Wisniewski and N. K. Robakis (2002). "A presenilin-1/γ-secretase cleavage releases the E-cadherin intracellular domain and regulates disassembly of adherens junctions." The EMBO Journal **21**(8): 1948-1956.

Martinelli, P. and E. I. Rugarli (2010). "Emerging roles of mitochondrial proteases in neurodegeneration." Biochimica et Biophysica Acta (BBA) - Bioenergetics **1797**(1): 1-10.

Marzi, S., T. Ribarska, A. R. Smith, E. Hannon, J. Poschmann, K. Moore, C. Troakes, S. Al-Sarraj, S. Beck, S. Newman, K. Lunnon, L. Schalkwyk and J. Mill (2017). "A histone acetylome-wide association study of Alzheimer's disease: neuropathology-associated regulatory variation in the human entorhinal cortex." bioRxiv.

Marzi, S. J. R., Teodora; Smith, Adam R; Hannon, Eilis; Poschmann, Jeremie;, K. T. Moore, Claire; Al-Sarraj, Safa; Beck, Stephan; Newman, Stuart; and K. S. Lunnon, Leonard C.; Mill, Jonathan (2018). "A histone acetylome-wide association study of Alzheimer's disease: neuropathologyassociated regulatory variation in the human entorhinal cortex." Submitted.

Masliah, E., M. Mallory, M. Alford, R. DeTeresa, L. A. Hansen, D. W. McKeel, Jr. and J. C. Morris (2001). "Altered expression of synaptic proteins occurs early during progression of Alzheimer's disease." Neurology **56**(1): 127-129.

Mastroeni, D., A. Grover, E. Delvaux, C. Whiteside, P. D. Coleman and J. Rogers (2010). "Epigenetic changes in Alzheimer's disease: decrements in DNA methylation." Neurobiol Aging **31**(12): 2025-2037.

Mastroeni, D., A. McKee, A. Grover, J. Rogers and P. D. Coleman (2009). "Epigenetic differences in cortical neurons from a pair of monozygotic twins discordant for Alzheimer's disease." PLoS One **4**(8): e6617.

Mastroeni, D., S. Sekar, J. Nolz, E. Delvaux, K. Lunnon, J. Mill, W. S. Liang and P. D. Coleman (2017). "ANK1 is up-regulated in laser captured microglia in Alzheimer's brain; the importance of addressing cellular heterogeneity." PLOS ONE **12**(7): e0177814.

Matarese, F., E. Carrillo-de Santa Pau and H. G. Stunnenberg (2011). "5-Hydroxymethylcytosine: a new kid on the epigenetic block?" Molecular Systems Biology **7**: 562-562.

Mattout, A. and E. Meshorer (2010). "Chromatin plasticity and genome organization in pluripotent embryonic stem cells." Current Opinion in Cell Biology **22**(3): 334-341.

Mattson, M. P. (2004). "Pathways towards and away from Alzheimer's disease." Nature **430**(7000): 631-639.

McIlwain, D. R., P. A. Lang, T. Maretzky, K. Hamada, K. Ohishi, S. K. Maney, T. Berger, A. Murthy, G. Duncan, H. C. Xu, K. S. Lang, D. Haussinger, A. Wakeham, A. Itie-Youten, R. Khokha, P. S. Ohashi, C. P. Blobel and T. W. Mak (2012). "iRhom2 regulation of TACE controls TNF-mediated protection against Listeria and responses to LPS." Science **335**(6065): 229-232.

McKeith, I. (2007). Dementia with Lewy bodies. Handbook of Clinical Neurology. W. C. Koller and E. Melamed, Elsevier. **84:** 531-548.

McKeith, I. G. (2002). "Dementia with Lewy bodies." The British Journal of Psychiatry **180**(2): 144-147.

McLaren, E. (2015). England and Wales; Mortality Statistics: Deaths Registered in England and Wales (Series DR). Scotland: National Records of Scotland Vital Events Reference Tables. Northern Ireland; Northern Ireland Statistic & Research Agency Registrar General Annual Report 2015. Office of National Statistics.

Meda, S. A., B. Narayanan, J. Liu, N. I. Perrone-Bizzozero, M. C. Stevens, V. D. Calhoun, D. C. Glahn, L. Shen, S. L. Risacher, A. J. Saykin and G. D. Pearlson (2012). "A large scale multivariate parallel ICA method reveals novel imaging-genetic relationships for Alzheimer's disease in the ADNI cohort." Neuroimage **60**(3): 1608-1621.

Meyniel, C. and P. Damier (2007). "Lewy body dementia and Parkinson disease dementia." Presse medicale (Paris, France : 1983) **36**(10 Pt 2): 1485-1490.

Mill, J. (2011). "Toward an integrated genetic and epigenetic approach to Alzheimer's disease." Neurobiol Aging **32**(7): 1188-1191.

Miraglia del Giudice, E., M. Francese, B. Nobili, L. Morle, S. Cutillo, J. Delaunay and S. Perrotta (1998). "High frequency of de novo mutations in ankyrin gene (ANK1) in children with hereditary spherocytosis." J Pediatr **132**(1): 117-120.

Miyashita, A., A. Koike, G. Jun, L. S. Wang, S. Takahashi, E. Matsubara, T. Kawarabayashi, M. Shoji, N. Tomita, H. Arai, T. Asada, Y. Harigaya, M. Ikeda, M. Amari, H. Hanyu, S. Higuchi, T. Ikeuchi, M. Nishizawa, M. Suga, Y. Kawase, H. Akatsu, K. Kosaka, T. Yamamoto, M. Imagawa, T. Hamaguchi, M. Yamada, T. Morihara, M. Takeda, T. Takao, K. Nakata, Y. Fujisawa, K. Sasaki, K. Watanabe, K. Nakashima, K. Urakami, T. Ooya, M. Takahashi, T. Yuzuriha, K. Serikawa, S. Yoshimoto, R. Nakagawa, J. W. Kim, C. S. Ki, H. H. Won, D. L. Na, S. W. Seo, I. Mook-Jung, C. Alzheimer Disease Genetics, P. St George-Hyslop, R. Mayeux, J. L. Haines, M. A. Pericak-Vance, M. Yoshida, N. Nishida, K. Tokunaga, K. Yamamoto, S. Tsuji, I. Kanazawa, Y. Ihara, G. D. Schellenberg, L. A. Farrer and R. Kuwano (2013). "SORL1 is genetically associated with late-onset Alzheimer's disease in Japanese, Koreans and Caucasians." PLoS One **8**(4): e58618.

Monson, N. L., S. J. Ireland, A. J. Ligocki, D. Chen, W. H. Rounds, M. Li, R. M. Huebinger, C. Munro Cullum, B. M. Greenberg, A. M. Stowe and R. Zhang (2014). "Elevated CNS inflammation in patients with preclinical Alzheimer's disease." J Cereb Blood Flow Metab **34**(1): 30-33.

Munzel, M., D. Globisch, T. Bruckl, M. Wagner, V. Welzmiller, S. Michalakis, M. Muller, M. Biel and T. Carell (2010). "Quantification of the sixth DNA base hydroxymethylcytosine in the brain." Angew Chem Int Ed Engl **49**(31): 5375-5377.

Myers, R. H. (2004). "Huntington's Disease Genetics." NeuroRx **1**(2): 255-262.

N'Diaye, E. N., C. S. Branda, S. S. Branda, L. Nevarez, M. Colonna, C. Lowell, J. A. Hamerman and W. E. Seaman (2009). "TREM-2 (triggering receptor expressed on myeloid cells 2) is a phagocytic receptor for bacteria." J Cell Biol **184**(2): 215-223.

Naj, A. C., G. Jun, G. W. Beecham, L. S. Wang, B. N. Vardarajan, J. Buros, P. J. Gallins, J. D. Buxbaum, G. P. Jarvik, P. K. Crane, E. B. Larson, T. D. Bird, B. F. Boeve, N. R. Graff-Radford, P. L. De Jager, D. Evans, J. A. Schneider, M. M. Carrasquillo, N. Ertekin-Taner, S. G. Younkin, C. Cruchaga, J. S. Kauwe, P. Nowotny, P. Kramer, J. Hardy, M. J. Huentelman, A. J. Myers, M. M. Barmada, F. Y. Demirci, C. T. Baldwin, R. C. Green, E. Rogaeva, P. St George-Hyslop, S. E.

Arnold, R. Barber, T. Beach, E. H. Bigio, J. D. Bowen, A. Boxer, J. R. Burke, N. J. Cairns, C. S. Carlson, R. M. Carney, S. L. Carroll, H. C. Chui, D. G. Clark, J. Corneveaux, C. W. Cotman, J. L. Cummings, C. DeCarli, S. T. DeKosky, R. Diaz-Arrastia, M. Dick, D. W. Dickson, W. G. Ellis, K. M. Faber, K. B. Fallon, M. R. Farlow, S. Ferris, M. P. Frosch, D. R. Galasko, M. Ganguli, M. Gearing, D. H. Geschwind, B. Ghetti, J. R. Gilbert, S. Gilman, B. Giordani, J. D. Glass, J. H. Growdon, R. L. Hamilton, L. E. Harrell, E. Head, L. S. Honig, C. M. Hulette, B. T. Hyman, G. A. Jicha, L. W. Jin, N. Johnson, J. Karlawish, A. Karydas, J. A. Kaye, R. Kim, E. H. Koo, N. W. Kowall, J. J. Lah, A. I. Levey, A. P. Lieberman, O. L. Lopez, W. J. Mack, D. C. Marson, F. Martiniuk, D. C. Mash, E. Masliah, W. C. McCormick, S. M. McCurry, A. N. McDavid, A. C. McKee, M. Mesulam, B. L. Miller, C. A. Miller, J. W. Miller, J. E. Parisi, D. P. Perl, E. Peskind, R. C. Petersen, W. W. Poon, J. F. Quinn, R. A. Rajbhandary, M. Raskind, B. Reisberg, J. M. Ringman, E. D. Roberson, R. N. Rosenberg, M. Sano, L. S. Schneider, W. Seeley, M. L. Shelanski, M. A. Slifer, C. D. Smith, J. A. Sonnen, S. Spina, R. A. Stern, R. E. Tanzi, J. Q. Trojanowski, J. C. Troncoso, V. M. Van Deerlin, H. V. Vinters, J. P. Vonsattel, S. Weintraub, K. A. Welsh-Bohmer, J. Williamson, R. L. Woltjer, L. B. Cantwell, B. A. Dombroski, D. Beekly, K. L. Lunetta, E. R. Martin, M. I. Kamboh, A. J. Saykin, E. M. Reiman, D. A. Bennett, J. C. Morris, T. J. Montine, A. M. Goate, D. Blacker, D. W. Tsuang, H. Hakonarson, W. A. Kukull, T. M. Foroud, J. L. Haines, R. Mayeux, M. A. Pericak-Vance, L. A. Farrer and G. D. Schellenberg (2011). "Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease." Nat Genet **43**(5): 436-441.

Nan, X., H. H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, R. N. Eisenman and A. Bird (1998). "Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex." Nature **393**(6683): 386-389.

Naslund, J., V. Haroutunian, R. Mohs, K. L. Davis, P. Davies, P. Greengard and J. D. Buxbaum (2000). "Correlation between elevated levels of amyloid beta-peptide in the brain and cognitive decline." Jama **283**(12): 1571-1577.

NHS. (2017). "Vascular Dementia." Retrieved 22 Jan 2018.

Norden, D. M. and J. P. Godbout (2013). "Review: microglia of the aged brain: primed to be activated and resistant to regulation." Neuropathol Appl Neurobiol **39**(1): 19-34.

Nyren, P. (1987). "Enzymatic method for continuous monitoring of DNA polymerase activity." Analytical Biochemistry **167**(2): 235-238.

Ooi, S. K., C. Qiu, E. Bernstein, K. Li, D. Jia, Z. Yang, H. Erdjument-Bromage, P. Tempst, S. P. Lin, C. D. Allis, X. Cheng and T. H. Bestor (2007). "DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA." Nature **448**(7154): 714-717.

Orlando, V. (2000). "Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation." Trends in Biochemical Sciences **25**(3): 99-104.

Otani, J., T. Nankumo, K. Arita, S. Inamoto, M. Ariyoshi and M. Shirakawa (2009). "Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain." EMBO Rep **10**(11): 1235-1241.

Pan, G., S. Tian, J. Nie, C. Yang, V. Ruotti, H. Wei, G. A. Jonsdottir, R. Stewart and J. A. Thomson (2007). "Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells." Cell Stem Cell **1**(3): 299-312.

Pan, X. D., Y. G. Zhu, N. Lin, J. Zhang, Q. Y. Ye, H. P. Huang and X. C. Chen (2011). "Microglial phagocytosis induced by fibrillar beta-amyloid is attenuated by oligomeric beta-amyloid: implications for Alzheimer's disease." Mol Neurodegener **6**: 45.

Pastor, W. A., U. J. Pape, Y. Huang, H. R. Henderson, R. Lister, M. Ko, E. M. McLoughlin, Y. Brudno, S. Mahapatra, P. Kapranov, M. Tahiliani, G. Q. Daley, X. S. Liu, J. R. Ecker, P. M. Milos, S. Agarwal and A. Rao (2011). "Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells." Nature **473**(7347): 394-397.

Pedersen, B. S., D. A. Schwartz, I. V. Yang and K. J. Kechris (2012). "Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values." Bioinformatics **28**(22): 2986-2988.

Perry, V. H. and C. Holmes (2014). "Microglial priming in neurodegenerative disease." Nat Rev Neurol **10**(4): 217-224.

Phipson, B., J. Maksimovic and A. Oshlack (2016). "missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform." Bioinformatics **32**(2): 286-288.

Pidsley, R., Y. W. CC, M. Volta, K. Lunnon, J. Mill and L. C. Schalkwyk (2013). "A data-driven approach to preprocessing Illumina 450K methylation array data." BMC Genomics **14**: 293.

Pidsley, R., C. C. Y Wong, M. Volta, K. Lunnon, J. Mill and L. C. Schalkwyk (2013). "A data-driven approach to preprocessing Illumina 450K methylation array data." BMC Genomics **14**(1): 1-10.

Pidsley, R., E. Zotenko, T. J. Peters, M. G. Lawrence, G. P. Risbridger, P. Molloy, S. Van Djik, B. Muhlhausler, C. Stirzaker and S. J. Clark (2016). "Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling." Genome Biol **17**(1): 208.

Pigott, K., J. Rick, S. X. Xie, H. Hurtig, A. Chen-Plotkin, J. E. Duda, J. F. Morley, L. M. Chahine, N. Dahodwala, R. S. Akhtar, A. Siderowf, J. Q. Trojanowski and D. Weintraub (2015). "Longitudinal study of normal cognition in Parkinson disease." Neurology **85**(15): 1276-1282.

Popa, G. (2017). DNA and Protein Synthesis: Histones.

Potkin, S. G., G. Guffanti, A. Lakatos, J. A. Turner, F. Kruggel, J. H. Fallon, A. J. Saykin, A. Orro, S. Lupoli, E. Salvi, M. Weiner and F. Macciardi (2009). "Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease." PLoS One **4**(8): e6501.

Price, E. M., A. M. Cotton, L. L. Lam, P. Farré, E. Emberly, C. J. Brown, W. P. Robinson and M. S. Kobor (2013). "Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array." Epigenetics & Chromatin **6**: 4-4.

Prince, M., W. Anders, G. Maëlenn, A. Gemma-Claire, W. Yu-Tzu and P. Matthew (2015). "World Alzheimer Report 2015 - The Global Impact of Dementia, an analysis of prevalence,incidence,cost and trends." Alzheimer's Disease International: 1-87.

Prince, M., M. Knapp, M. Guerchet, P. McCrone, M. Prina, A. Comas-Herrera, R. Wittenberg, B. Adelaja, B. Hu, D. King, A. Rehill and D. Salimkumar (2014). Dementia UK: Update.

Prince, M., M. Knapp, M. Guerchet, P. McCrone, M. Prina, A. Comas-Herrera, R. Wittenberg, B. Adelaja, B. Hu, D. King, A. Rehill and D. Salimkumar (2014). Dementia UK: Update Second Edition report produced by King's College London and the London School of Economics for the Alzheimer's Societ.

Prince, M., A. Wimo, M. Guerchet, G. Ali, Y. Wu and M. Prina (2015). World Alzheimer's Report 2015, The global prevalence of dementia: An analysis of prevalence, incidence, cost and trends Alzheimers Dement, Alzheimer's Disease International

Profenno, L. A., A. P. Porsteinsson and S. V. Faraone (2010). "Meta-analysis of Alzheimer's disease risk with obesity, diabetes, and related disorders." Biol Psychiatry **67**(6): 505-512.

Puglielli, L., R. E. Tanzi and D. M. Kovacs (2003). "Alzheimer's disease: the cholesterol connection." Nat Neurosci **6**(4): 345-351.

Querfurth, H. W. and F. M. LaFerla (2010). "Alzheimer's Disease." New England Journal of Medicine **362**(4): 329-344.

R Development Core Team (2012). "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna, Austria 2012.

Ramanan, V. K., S. L. Risacher, K. Nho, S. Kim, S. Swaminathan, L. Shen, T. M. Foroud, H. Hakonarson, M. J. Huentelman, P. S. Aisen, R. C. Petersen, R. C. Green, C. R. Jack, R. A. Koeppe, W. J. Jagust, M. W. Weiner and A. J. Saykin (2014). "APOE and BCHE as modulators of cerebral

amyloid deposition: a florbetapir PET genome-wide association study." Mol Psychiatry **19**(3): 351-357.

Randon, J., E. Miraglia del Giudice, M. Bozon, S. Perrotta, M. De Vivo, A. Iolascon, J. Delaunay and L. Morle (1997). "Frequent de novo mutations of the ANK1 gene mimic a recessive mode of transmission in hereditary spherocytosis: three new ANK1 variants: ankyrins Bari, Napoli II and Anzio." Br J Haematol **96**(3): 500-506.

Rao, J. S., V. L. Keleshian, S. Klein and S. I. Rapoport (2012). "Epigenetic modifications in frontal cortex from Alzheimer's disease and bipolar disorder patients." Transl Psychiatry **2**: e132.

Rauch, T. A., Z. Wang, X. Wu, K. H. Kernstine, A. D. Riggs and G. P. Pfeifer (2012). "DNA methylation biomarkers for lung cancer." Tumour Biol **33**(2): 287-296.

Rauch, T. A., X. Wu, X. Zhong, A. D. Riggs and G. P. Pfeifer (2009). "A human B cell methylome at 100-base pair resolution." Proc Natl Acad Sci U S A **106**(3): 671-678.

Reiman, E. M., J. A. Webster, A. J. Myers, J. Hardy, T. Dunckley, V. L. Zismann, K. D. Joshipura, J. V. Pearson, D. Hu-Lince, M. J. Huentelman, D. W. Craig, K. D. Coon, W. S. Liang, R. H. Herbert, T. Beach, K. C. Rohrer, A. S. Zhao, D. Leung, L. Bryden, L. Marlowe, M. Kaleem, D. Mastroeni, A. Grover, C. B. Heward, R. Ravid, J. Rogers, M. L. Hutton, S. Melquist, R. C. Petersen, G. E. Alexander, R. J. Caselli, W. Kukull, A. Papassotiropoulos and D. A. Stephan (2007). "GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers." Neuron **54**(5): 713-720.

Reiner, A., I. Dragatsis and P. Dietrich (2011). "GENETICS AND NEUROPATHOLOGY OF HUNTINGTON'S DISEASE." International review of neurobiology **98**: 325-372.

Relton, C. L. and G. Davey Smith (2012). "Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease." Int J Epidemiol **41**(1): 161-176.

Rezeli, M., H. Zetterberg, K. Blennow, A. Brinkmalm, T. Laurell, O. Hansson and G. Marko-Varga (2015). "Quantification of total apolipoprotein E and its specific isoforms in cerebrospinal fluid and blood in Alzheimer's disease and other neurodegenerative diseases." EuPA Open Proteomics **8**: 137-143.

Rhoads, A. and K. F. Au (2015). "PacBio Sequencing and Its Applications." Genomics, Proteomics & Bioinformatics **13**(5): 278-289.

Ridge, P. G., S. Mukherjee, P. K. Crane, J. S. Kauwe and C. Alzheimer's Disease Genetics (2013). "Alzheimer's disease: analyzing the missing heritability." PLoS One **8**(11): e79771.

Rogaeva, E., Y. Meng, J. H. Lee, Y. Gu, T. Kawarai, F. Zou, T. Katayama, C. T. Baldwin, R. Cheng, H. Hasegawa, F. Chen, N. Shibata, K. L. Lunetta, R. Pardossi-Piquard, C. Bohm, Y. Wakutani, L. A. Cupples, K. T. Cuenco, R. C. Green, L. Pinessi, I. Rainero, S. Sorbi, A. Bruni, R. Duara, R. P. Friedland, R. Inzelberg, W. Hampe, H. Bujo, Y. Song, O. Andersen, T. E. Willnow, N. Graff-Radford, R. Petersen, D. Dickson, S. D. Der, P. E. Fraser, G. Schmitt-Ulms, S. Younkin, R. Mayeux, L. A. Farrer and P. St George-Hyslop (2007). "The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer's Disease." Nature genetics **39**(2): 168-177.

Román, G. C., T. K. Tatemichi, T. Erkinjuntti, J. L. Cummings, J. C. Masdeu, J. H. Garcia, L. Amaducci, J.-M. Orgogozo, A. Brun, A. Hofman, D. M. Moody, M. D. O'Brien, T. Yamaguchi, J. Grafman, B. P. Drayer, D. A. Bennett, M. Fisher, J. Ogata, E. Kokmen, F. Bermejo, P. A. Wolf, P. B. Gorelick, K. L. Bick, A. K. Pajeau, M. A. Bell, C. DeCarli, A. Culebras, A. D. Korczyn, J. Bogousslavsky, A. Hartmann and P. Scheinberg (1993). "Vascular dementia." Diagnostic criteria for research studies: Report of the NINDS-AIREN International Workshop* **43**(2): 250-250.

Roos, R. A. C. (2010). "Huntington's disease: a clinical review." Orphanet Journal of Rare Diseases **5**: 40-40.

Rose, N. R. and R. J. Klose (2014). "Understanding the relationship between DNA methylation and histone lysine methylation." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1839**(12): 1362-1372.

Rosenberg, C. K., T. J. Cummings, A. M. Saunders, C. Widico, L. M. McIntyre and C. M. Hulette (2001). "Dementia with Lewy bodies and Alzheimer's disease." Acta Neuropathologica **102**(6): 621-626.

Roses, A. D., M. W. Lutz, M. J. Huentelman, O. Chiba-Falek, K. A. Welsh-Bohmer and E. M. Reiman (2009). "Apoeε3 A3nd Tomm-40 Haplotypes Determine Inheritance of Alzheimer's Disease Independently of Apoeε4 Risk." Alzheimer's & Dementia: The Journal of the Alzheimer's Association **5**(4): e1.

Roses, M. D., Allen D. (1996). "APOLIPOPROTEIN E ALLELES AS RISK FACTORS IN ALZHEIMER'S DISEASE." Annual Review of Medicine **47**(1): 387-400.

Rouch, L., P. Cestac, O. Hanon, C. Cool, C. Helmer, B. Bouhanick, B. Chamontin, J. F. Dartigues, B. Vellas and S. Andrieu (2015). "Antihypertensive drugs, prevention of cognitive decline and dementia: a systematic review of observational studies, randomized controlled trials and meta-analyses, with discussion of potential mechanisms." CNS Drugs **29**(2): 113-130.

Sá, F., P. Pinto, C. Cunha, R. Lemos, L. Letra, M. Simões and I. Santana (2012). "Differences between Early and Late-Onset Alzheimer's Disease in Neuropsychological Tests." Frontiers in Neurology **3**: 81.

Sadakierska-Chudy, A. and M. Filip (2015). "A Comprehensive View of the Epigenetic Landscape. Part II: Histone Post-translational Modification, Nucleosome Level, and Chromatin Regulation by ncRNAs." Neurotoxicity Research **27**: 172-197.

Sakamuro, D., K. J. Elliott, R. Wechsler-Reya and G. C. Prendergast (1996). "BIN1 is a novel MYC-interacting protein with features of a tumour suppressor." Nat Genet **14**(1): 69-77.

Salminen, A., J. Ojala, A. Kauppinen, K. Kaarniranta and T. Suuronen (2009). "Inflammation in Alzheimer's disease: Amyloid-β oligomers trigger innate immunity defence via pattern recognition receptors." Progress in Neurobiology **87**(3): 181-194.

Sambrook, J. and D. W. Russell (2006). "Purification of Nucleic Acids by Extraction with Phenol:Chloroform." Cold Spring Harbor Protocols **2006**(1): pdb.prot4455.

Sampson, E. L., M. R. Blanchard, L. Jones, A. Tookman and M. King (2009). "Dementia in the acute hospital: prospective cohort study of prevalence and mortality." Br J Psychiatry **195**(1): 61-66.

Sanan, D. A., K. H. Weisgraber, S. J. Russell, R. W. Mahley, D. Huang, A. Saunders, D. Schmechel, T. Wisniewski, B. Frangione and A. D. Roses (1994). "Apolipoprotein E associates with beta amyloid peptide of Alzheimer's disease to form novel monofibrils. Isoform apoE4 associates more efficiently than apoE3." Journal of Clinical Investigation **94**(2): 860-869.

Sanchez-Mut, J. V., E. Aso, H. Heyn, T. Matsuda, C. Bock, I. Ferrer and M. Esteller (2014). "Promoter Hypermethylation of the Phosphatase DUSP22 Mediates PKA-Dependent TAU Phosphorylation and CREB Activation in Alzheimer's Disease." Hippocampus **24**(4): 363-368.

Sanchez-Mut, J. V., E. Aso, H. Heyn, T. Matsuda, C. Bock, I. Ferrer and M. Esteller (2014). "Promoter hypermethylation of the phosphatase DUSP22 mediates PKA-dependent TAU phosphorylation and CREB activation in Alzheimer's disease." Hippocampus **24**(4): 363-368.

Sanchez-Mut, J. V., E. Aso, N. Panayotis, I. Lott, M. Dierssen, A. Rabano, R. G. Urdinguio, A. F. Fernandez, A. Astudillo, J. I. Martin-Subero, B. Balint, M. F. Fraga, A. Gomez, C. Gurnot, J. C. Roux, J. Avila, T. K. Hensch, I. Ferrer and M. Esteller (2013). "DNA methylation map of mouse and human brain identifies target genes in Alzheimer's disease." Brain **136**(Pt 10): 3018-3027.

Sanchez-Mut, J. V., H. Heyn, E. Vidal, S. Moran, S. Sayols, R. Delgado-Morales, M. D. Schultz, B. Ansoleaga, P. Garcia-Esparcia, M. Pons-Espinal, M. M. de Lagran, J. Dopazo, A. Rabano, J. Avila, M. Dierssen, I. Lott, I. Ferrer, J. R. Ecker and M. Esteller (2016). "Human DNA methylomes of neurodegenerative diseases show common epigenomic patterns." Transl Psychiatry **6**: e718.

Sander, J. D. and J. K. Joung (2014). "CRISPR-Cas systems for editing, regulating and targeting genomes." Nat Biotechnol **32**(4): 347-355.

Santos-Rosa, H., R. Schneider, A. J. Bannister, J. Sherriff, B. E. Bernstein, N. C. T. Emre, S. L. Schreiber, J. Mellor and T. Kouzarides (2002). "Active genes are tri-methylated at K4 of histone H3." Nature **419**: 407.

Schadt, E. E., S. Turner and A. Kasarskis (2010). "A window into third-generation sequencing." Hum Mol Genet **19**(R2): R227-240.

Scheuner, D., C. Eckman, M. Jensen, X. Song, M. Citron, N. Suzuki, T. D. Bird, J. Hardy, M. Hutton, W. Kukull, E. Larson, E. Levy-Lahad, M. Viitanen, E. Peskind, P. Poorkaj, G. Schellenberg, R. Tanzi, W. Wasco, L. Lannfelt, D. Selkoe and S. Younkin (1996). "Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease." Nat Med **2**(8): 864-870.

Schübeler, D., D. M. MacAlpine, D. Scalzo, C. Wirbelauer, C. Kooperberg, F. van Leeuwen, D. E. Gottschling, L. P. O'Neill, B. M. Turner, J. Delrow, S. P. Bell and M. Groudine (2004). "The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote." Genes & Development **18**(11): 1263-1271.

Seshadri, S., A. L. Fitzpatrick, M. A. Ikram, A. L. DeStefano, V. Gudnason, M. Boada, J. C. Bis, A. V. Smith, M. M. Carassquillo, J. C. Lambert, D. Harold, E. M. Schrijvers, R. Ramirez-Lorca, S. Debette, W. T. Longstreth, Jr., A. C. Janssens, V. S. Pankratz, J. F. Dartigues, P. Hollingworth, T. Aspelund, I. Hernandez, A. Beiser, L. H. Kuller, P. J. Koudstaal, D. W. Dickson, C. Tzourio, R. Abraham, C. Antunez, Y. Du, J. I. Rotter, Y. S. Aulchenko, T. B. Harris, R. C. Petersen, C. Berr, M. J. Owen, J. Lopez-Arrieta, B. N. Varadarajan, J. T. Becker, F. Rivadeneira, M. A. Nalls, N. R. Graff-Radford, D. Campion, S. Auerbach, K. Rice, A. Hofman, P. V. Jonsson, H. Schmidt, M. Lathrop, T. H. Mosley, R. Au, B. M. Psaty, A. G. Uitterlinden, L. A. Farrer, T. Lumley, A. Ruiz, J. Williams, P. Amouyel, S. G. Younkin, P. A. Wolf, L. J. Launer, O. L. Lopez, C. M. van Duijn and M. M. Breteler (2010). "Genome-wide analysis of genetic loci associated with Alzheimer disease." JAMA **303**(18): 1832-1840.

Sheng, J. G., R. E. Mrak and W. S. Griffin (1997). "Neuritic plaque evolution in Alzheimer's disease is accompanied by transition of activated microglia from primed to enlarged to phagocytic forms." Acta Neuropathol **94**(1): 1-5.

Sherwani, S. I. and H. A. Khan (2015). "Role of 5-hydroxymethylcytosine in neurodegeneration." Gene **570**(1): 17-24.

Siggs, O. M., N. Xiao, Y. Wang, H. Shi, W. Tomisato, X. Li, Y. Xia and B. Beutler (2012). "iRhom2 is required for the secretion of mouse TNFalpha." Blood **119**(24): 5769-5771.

Silva, P. N., T. K. Furuya, I. L. Braga, L. T. Rasmussen, R. W. Labio, P. H. Bertolucci, E. S. Chen, G. Turecki, N. Mechawar, S. L. Payao, J. Mill and M. C. Smith (2014). "Analysis of HSPA8 and HSPA9 mRNA expression and promoter methylation in the brain and blood of Alzheimer's disease patients." J Alzheimers Dis **38**(1): 165-170.

Simon, R., M. Girod, C. Fonbonne, A. Salvador, Y. Clément, P. Lantéri, P. Amouyel, J. C. Lambert and J. Lemoine (2012). "Total ApoE and ApoE4 Isoform Assays in an Alzheimer's Disease Case-control Study by Targeted Mass Spectrometry (n = 669): A Pilot Assay for Methionine-containing Proteotypic Peptides." Molecular & Cellular Proteomics **11**(11): 1389-1403.

Skvortsova, K., E. Zotenko, P.-L. Luu, C. M. Gould, S. S. Nair, S. J. Clark and C. Stirzaker (2017). "Comprehensive evaluation of genome-wide 5-hydroxymethylcytosine profiling approaches in human DNA." Epigenetics & Chromatin **10**: 16.

Sleegers, K., J. C. Lambert, L. Bertram, M. Cruts, P. Amouyel and C. Van Broeckhoven (2010). "The pursuit of susceptibility genes for Alzheimer's disease: progress and prospects." Trends Genet **26**(2): 84-93.

Slieker, R. C., S. D. Bos, J. J. Goeman, J. V. Bovee, R. P. Talens, R. van der Breggen, H. E. Suchiman, E. W. Lameijer, H. Putter, E. B. van den Akker, Y. Zhang, J. W. Jukema, P. E. Slagboom, I. Meulenbelt and B. T. Heijmans (2013). "Identification and systematic annotation of tissue-

specific differentially methylated regions using the Illumina 450k array." Epigenetics Chromatin **6**(1): 26.
Slooter, A. J., M. Cruts, S. Kalmijn, A. Hofman, M. M. Breteler, C. Van Broeckhoven and C. M. van Duijn (1998). "Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study." Arch Neurol **55**(7): 964-968.
Smith, A. R., J. Mill, R. G. Smith and K. Lunnon (2016). "Elucidating novel dysfunctional pathways in Alzheimer's disease by integrating loci identified in genetic and epigenetic studies." Neuroepigenetics **6**: 32-50.
Smith, A. R., R. G. Smith, J. Burrage, C. Troakes, S. Al-Sarraj, R. N. Kalaria, C. Sloan, A. C. Robinson, J. Mill and K. Lunnon (2018). "A cross-brain-regions study of *ANK1* DNA methylation in different neurodegenerative diseases. ." Submitted.
Smith, A. R., R. G. Smith, D. Condliffe, E. Hannon, L. Schalkwyk, J. Mill and K. Lunnon (2016). "Increased DNA methylation near TREM2 is consistently seen in the superior temporal gyrus in Alzheimer's disease brain." Neurobiol Aging **47**: 35-40.
Smith, A. R., R. G. Smith, E. Hannon and J. A. Y. B. Roubroeks, Joe. Troakes, Claire. Al-Sarraj, Safa. Mill, Jonathan. van den Hove, Daniel L.. Lunnon, Katie (2018). "Parallel profiling of DNA methylation and hydroxymethylation highlights neuropathology-associated epigenetic variation within ANK1 and thirteen novel loci in Alzheimer's disease entorhinal cortex." Submitted.
Smith, A. R. L., Katie (2018). Post-Mortem Brain Tissue And Its Application To Study Epigenetic Regulation In Alzheimer's Disease. Post Mortem Neurochemistry. V. Balcar. Springer Nature. **Under Review**
Smith, A. R. S., Rebecca G; Hannon, Eilis; Roubroeks, Janou A Y; Burrage, Joe; Troakes, Claire; Al-Sarraj, Safa; Mill, Jonathan; van den Hove, Daniel L; Lunnon Katie (2018). "Parallel profiling of DNA methylation and hydroxymethylation highlights neuropathology-associated epigenetic variation within ANK1 and thirteen novel loci in Alzheimer's disease entorhinal cortex." Submitted.
Smith, R. G., E. Hannon, P. De Jager, L. Chibnik, S. Lott, A. R. Smith, D. Condliffe and K. Lunnon (2018). "Elevated DNA methylation across a 48kb region spanning the HOXA gene cluster on chromosome 7 is associated with Alzheimer's disease neuropathology in the prefrontal cortex and superior temporal gyrus." Alzheimer's & Dementia.
Smith, R. G. and K. Lunnon (2017). DNA Modifiactions and Alzheimer's Disease. Neuroepigenomics in Aging and Disease. . R. Delgado-Morales, Springer International Publishing**:** 303-319.
Song, C. X., K. E. Szulwach, Y. Fu, Q. Dai, C. Yi and X. Li (2011). "Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine." Nat Biotechnol **29**.
Spiers, H., E. Hannon, L. C. Schalkwyk, N. J. Bray and J. Mill (2017). "5-hydroxymethylcytosine is highly dynamic across human fetal brain development." BMC Genomics **18**(1): 738.
Spillantini, M. G., M. L. Schmidt, V. M. Lee, J. Q. Trojanowski, R. Jakes and M. Goedert (1997). "Alpha-synuclein in Lewy bodies." Nature **388**(6645): 839-840.
Spitz, F. and E. E. Furlong (2012). "Transcription factors: from enhancer binding to developmental control." Nat Rev Genet **13**(9): 613-626.
Spruijt, C. G., F. Gnerlich, A. H. Smits, T. Pfaffeneder, P. W. Jansen, C. Bauer, M. Munzel, M. Wagner, M. Muller, F. Khan, H. C. Eberl, A. Mensinga, A. B. Brinkman, K. Lephikov, U. Muller, J. Walter, R. Boelens, H. van Ingen, H. Leonhardt, T. Carell and M. Vermeulen (2013). "Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives." Cell **152**(5): 1146-1159.
Stefanis, L. (2012). "α-Synuclein in Parkinson's Disease." Cold Spring Harbor Perspectives in Medicine **2**(2): a009399.
Steven Henikoff and M. A. Matzke (1997). "Exploring and explaining epigenetic effects." Editorial **13**(8).
Streit, W. J., N. W. Sammons, A. J. Kuhns and D. L. Sparks (2004). "Dystrophic microglia in the aging human brain." Glia **45**(2): 208-212.

Strittmatter, W. J., A. M. Saunders, D. Schmechel, M. Pericak-Vance, J. Enghild, G. S. Salvesen and A. D. Roses (1993). "Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease." Proceedings of the National Academy of Sciences of the United States of America **90**(5): 1977-1981.

Strittmatter, W. J., K. H. Weisgraber, D. Y. Huang, L. M. Dong, G. S. Salvesen, M. Pericak-Vance, D. Schmechel, A. M. Saunders, D. Goldgaber and A. D. Roses (1993). "Binding of human apolipoprotein E to synthetic amyloid beta peptide: isoform-specific effects and implications for late-onset Alzheimer disease." Proceedings of the National Academy of Sciences of the United States of America **90**(17): 8098-8102.

Sun-Chong Wang, B. O., Axel Schumacher (2008). "Age specific epigenetic drift in late onset Alzheimer's disease." PLoS ONE **3**(7).

Szulwach, K. E., X. Li, Y. Li, C.-X. Song, J. W. Han, S. Kim, S. Namburi, K. Hermetz, J. J. Kim, M. K. Rudd, Y.-S. Yoon, B. Ren, C. He and P. Jin (2011). "Integrating 5-Hydroxymethylcytosine into the Epigenomic Landscape of Human Embryonic Stem Cells." PLOS Genetics **7**(6): e1002154.

Szulwach, K. E., X. Li, Y. Li, C. X. Song, H. Wu, Q. Dai, H. Irier, A. K. Upadhyay, M. Gearing, A. I. Levey, A. Vasanthakumar, L. A. Godley, Q. Chang, X. Cheng, C. He and P. Jin (2011). "5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging." Nat Neurosci **14**(12): 1607-1616.

Szwagierczak, A., S. Bultmann, C. S. Schmidt, F. Spada and H. Leonhardt (2010). "Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA." Nucleic Acids Research **38**(19): e181-e181.

Taberlay, Phillippa C., Theresa K. Kelly, C.-C. Liu, Jueng S. You, Daniel D. De Carvalho, Tina B. Miranda, Xianghong J. Zhou, G. Liang and Peter A. Jones (2011). "Polycomb-Repressed Genes Have Permissive Enhancers that Initiate Reprogramming." Cell **147**(6): 1283-1294.

Tahiliani, M., K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala and Y. Brudno (2009). "Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1." Science **324**.

Takahashi, K., C. D. Rochford and H. Neumann (2005). "Clearance of apoptotic neurons without inflammation by microglial triggering receptor expressed on myeloid cells-2." J Exp Med **201**(4): 647-657.

Takei, N., A. Miyashita, T. Tsukie, H. Arai, T. Asada, M. Imagawa, M. Shoji, S. Higuchi, K. Urakami, H. Kimura, A. Kakita, H. Takahashi, S. Tsuji, I. Kanazawa, Y. Ihara, S. Odani and R. Kuwano (2009). "Genetic association study on in and around the APOE in late-onset Alzheimer disease in Japanese." Genomics **93**(5): 441-448.

Toledo, E. M. and N. C. Inestrosa (2010). "Activation of Wnt signaling by lithium and rosiglitazone reduced spatial memory impairment and neurodegeneration in brains of an APPswe/PSEN1DeltaE9 mouse model of Alzheimer's disease." Mol Psychiatry **15**(3): 272-285, 228.

Tsai, R. Y. (2015). "p53-guided response to nucleostemin loss in normal versus cancer cells." Cell Death Dis **6**: e2030.

Tuppo, E. E. and H. R. Arias (2005). "The role of inflammation in Alzheimer's disease." Int J Biochem Cell Biol **37**(2): 289-305.

Untergasser, A., H. Nijveen, X. Rao, T. Bisseling, R. Geurts and J. A. Leunissen (2007). "Primer3Plus, an enhanced web interface to Primer3." Nucleic Acids Res **35**(Web Server issue): W71-74.

Urdinguio, R. G., J. V. Sanchez-Mut and M. Esteller (2009). "Epigenetic mechanisms in neurological diseases: genes, syndromes, and therapies." The Lancet Neurology **8**(11): 1056-1072.

Van Den Berg, D. J., A. K. Sharma, E. Bruno and R. Hoffman (1998). "Role of Members of the Wnt Gene Family in Human Hematopoiesis." Blood **92**(9): 3189-3202.

Vienna, R. D. C. T. (2012). R Foundation for Statistical Computing

Villar-Menendez, I., M. Blanch, S. Tyebji, T. Pereira-Veiga, J. L. Albasanz, M. Martin, I. Ferrer, E. Perez-Navarro and M. Barrachina (2013). "Increased 5-methylcytosine and decreased 5-

hydroxymethylcytosine levels are associated with reduced striatal A2AR levels in Huntington's disease." Neuromolecular Med **15**(2): 295-309.
Voronin, D. A. and E. V. Kiseleva (2008). "Functional role of proteins containing ankyrin repeats." Cell and Tissue Biology **2**(1): 1-12.
Walker, F. O. (2007). "Huntington's disease." The Lancet **369**(9557): 218-228.
Walker, M. P., F. M. LaFerla, S. S. Oddo and G. J. Brewer (2013). "Reversible epigenetic histone modifications and Bdnf expression in neurons with aging and from a mouse model of Alzheimer's disease." AGE **35**(3): 519-531.
Wang, F., Y. Yang, X. Lin, J. Q. Wang, Y. S. Wu, W. Xie, D. Wang, S. Zhu, Y. Q. Liao, Q. Sun, Y. G. Yang, H. R. Luo, C. Guo, C. Han and T. S. Tang (2013). "Genome-wide loss of 5-hmC is a novel epigenetic feature of Huntington's disease." Hum Mol Genet **22**(18): 3641-3653.
Wang, T., W. Guan, J. Lin, N. Boutaoui, G. Canino, J. Luo, J. C. Celedon and W. Chen (2015). "A systematic study of normalization methods for Infinium 450K methylation data using whole-genome bisulfite sequencing data." Epigenetics **10**(7): 662-669.
Wang, T., Q. Pan, L. Lin, K. E. Szulwach, C.-X. Song, C. He, H. Wu, S. T. Warren, P. Jin, R. Duan and X. Li (2012). "Genome-wide DNA hydroxymethylation changes are associated with neurodevelopmental genes in the developing human cerebellum." Human Molecular Genetics **21**(26): 5500-5510.
Warner, T. T. and A. H. Schapira (2003). "Genetic and environmental factors in the cause of Parkinson's disease." Ann Neurol **53 Suppl 3**: S16-23; discussion S23-15.
Watson, C. T., P. Roussos, P. Garg, D. J. Ho, N. Azam, P. L. Katsel, V. Haroutunian and A. J. Sharp (2016). "Genome-wide DNA methylation profiling in the superior temporal gyrus reveals epigenetic signatures associated with Alzheimer's disease." Genome Med **8**(1): 5.
Watson, C. T., P. Roussos, P. Garg, D. J. Ho, N. Azam, P. L. Katsel, V. Haroutunian and A. J. Sharp (2016). "Genome-wide DNA methylation profiling in the superior temporal gyrus reveals epigenetic signatures associated with Alzheimer's disease." Genome Medicine **8**(1): 5.
Weber, M., I. Hellmann, M. B. Stadler, L. Ramos, S. Paabo, M. Rebhan and D. Schubeler (2007). "Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome." Nat Genet **39**(4): 457-466.
Wechsler-Reya, R., D. Sakamuro, J. Zhang, J. Duhadaway and G. C. Prendergast (1997). "Structural analysis of the human BIN1 gene. Evidence for tissue-specific transcriptional regulation and alternate RNA splicing." J Biol Chem **272**(50): 31453-31458.
West, R. L., J. M. Lee and L. E. Maroun (1995). "Hypomethylation of the amyloid precursor protein gene in the brain of an alzheimer's disease patient." Journal of Molecular Neuroscience **6**(2): 141-146.
Whitehouse, P., D. Price, R. Struble, A. Clark, J. Coyle and M. Delon (1982). "Alzheimer's disease and senile dementia: loss of neurons in the basal forebrain." Science **215**(4537): 1237-1239.
Wigge, P., K. Kohler, Y. Vallis, C. A. Doyle, D. Owen, S. P. Hunt and H. T. McMahon (1997). "Amphiphysin heterodimers: potential role in clathrin-mediated endocytosis." Mol Biol Cell **8**(10): 2003-2015.
Wigge, P. and H. T. McMahon (1998). "The amphiphysin family of proteins and their role in endocytosis at the synapse." Trends in Neurosciences **21**(8): 339-344.
Wijsman, E. M., N. D. Pankratz, Y. Choi, J. H. Rothstein, K. M. Faber, R. Cheng, J. H. Lee, T. D. Bird, D. A. Bennett, R. Diaz-Arrastia, A. M. Goate, M. Farlow, B. Ghetti, R. A. Sweet, T. M. Foroud and R. Mayeux (2011). "Genome-wide association of familial late-onset Alzheimer's disease replicates BIN1 and CLU and nominates CUGBP2 in interaction with APOE." PLoS Genet **7**(2): e1001308.
Wimo, A., M. Guerchet, G.-C. Ali, Y.-T. Wu, A. M. Prina, B. Winblad, L. Jönsson, Z. Liu and M. Prince (2017). "The worldwide costs of dementia 2015 and comparisons with 2010." Alzheimer's & Dementia **13**(1): 1-7.

Wu, H., A. C. D'Alessio, S. Ito, Z. Wang, K. Cui and K. Zhao (2011). "Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells." Genes Dev **25**.

Wyatt, G. R. and S. S. Cohen (1952). "A new pyrimidine base from bacteriophage nucleic acids." Nature **170**(4338): 1072-1073.

Xu, Y., F. Wu, L. Tan, L. Kong, L. Xiong, J. Deng, A. J. Barbera, L. Zheng, H. Zhang, S. Huang, J. Min, T. Nicholson, T. Chen, G. Xu, Y. Shi, K. Zhang and Y. G. Shi (2011). "Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells." Mol Cell **42**(4): 451-464.

Yang, M., W. Ge, R. Chowdhury, T. D. Claridge, H. B. Kramer, B. Schmierer, M. A. McDonough, L. Gong, B. M. Kessler, P. J. Ratcliffe, M. L. Coleman and C. J. Schofield (2011). "Asparagine and aspartate hydroxylation of the cytoskeletal ankyrin family is catalyzed by factor-inhibiting hypoxia-inducible factor." J Biol Chem **286**(9): 7648-7660.

Yang, S., T. Liu, S. Li, X. Zhang, Q. Ding, H. Que, X. Yan, K. Wei and S. Liu (2008). "Comparative proteomic analysis of brains of naturally aging mice." Neuroscience **154**(3): 1107-1120.

Yang, X., H. Han, Daniel D. De Carvalho, Fides D. Lay, Peter A. Jones and G. Liang (2014). "Gene Body Methylation Can Alter Gene Expression and Is a Therapeutic Target in Cancer." Cancer Cell **26**(4): 577-590.

Yildirim, O., R. Li, J. H. Hung, P. B. Chen, X. Dong and L. S. Ee (2011). "Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells." Cell **147**.

Yocum, A. O., L. A. Steiner, N. E. Seidel, A. P. Cline, E. D. Rout, J. Y. Lin, C. Wong, L. J. Garrett, P. G. Gallagher and D. M. Bodine (2012). "A tissue-specific chromatin loop activates the erythroid ankyrin-1 promoter." Blood **120**(17): 3586-3593.

Yu, C. E., H. Seltman, E. R. Peskind, N. Galloway, P. X. Zhou, E. Rosenthal, E. M. Wijsman, D. W. Tsuang, B. Devlin and G. D. Schellenberg (2007). "Comprehensive analysis of APOE and selected proximate markers for late-onset Alzheimer's disease: patterns of linkage disequilibrium and disease/marker association." Genomics **89**(6): 655-665.

Yu, L., L. B. Chibnik, G. P. Srivastava, N. Pochet, J. Yang, J. Xu, J. Kozubek, N. Obholzer, S. E. Leurgans and J. A. Schneider (2014). "Association of Brain DNA Methylation in SORL1, ABCA7, HLA-DRB5, SLC24A4, and BIN1 With Pathological Diagnosis of Alzheimer Disease." JAMA neurology.

Zentner, G. E. and S. Henikoff (2013). "Regulation of nucleosome dynamics by histone modifications." Nature Structural &Amp; Molecular Biology **20**: 259.

Zhang, H., J. K. Kim, C. A. Edwards, Z. Xu, R. Taichman and C. Y. Wang (2005). "Clusterin inhibits apoptosis by interacting with activated Bax." Nat Cell Biol **7**(9): 909-915.

Zhang, K., M. Schrag, A. Crofton, R. Trivedi, H. Vinters and W. Kirsch (2012). "Targeted proteomics for quantification of histone acetylation in Alzheimer's disease." PROTEOMICS **12**(8): 1261-1268.

Zhao, J., Y. Zhu, J. Yang, L. Li, H. Wu, P. L. De Jager, P. Jin and D. A. Bennett (2017). "A genome-wide profiling of brain DNA hydroxymethylation in Alzheimer's disease." Alzheimer's & Dementia.

Zhao, J., Y. Zhu, J. Yang, L. Li, H. Wu, P. L. De Jager, P. Jin and D. A. Bennett (2017). "A genome-wide profiling of brain DNA hydroxymethylation in Alzheimer's disease." Alzheimers Dement **13**(6): 674-688.

Zhubi, A., Y. Chen, E. Dong, E. H. Cook, A. Guidotti and D. R. Grayson (2014). "Increased binding of MeCP2 to the GAD1 and RELN promoters may be mediated by an enrichment of 5-hmC in autism spectrum disorder (ASD) cerebellum." Transl Psychiatry **4**: e349.